Minimum Wages, Efficiency and Welfare*

David Berger   Kyle Herkenhoff   Simon Mongey

June 2, 2025

Abstract

Many argue that minimum wages can prevent efficiency losses from monopsony power. We assess this argument in a general equilibrium model of oligopsonistic labor markets with heterogeneous workers and firms. We decompose welfare gains into an *efficiency* component that captures reductions in monopsony power and a *redistributive* component that captures the way minimum wages shift resources across people. The minimum wage that maximizes the efficiency component of welfare lies below $8.00 and yields gains worth less than 0.2% of lifetime consumption. When we add back in Utilitarian redistributive motives, the optimal minimum wage is $11 and redistribution accounts for 102.5% of the resulting welfare gains, implying offsetting efficiency losses of -2.5%. The reason a minimum wage struggles to deliver efficiency gains is that with realistic firm productivity dispersion, a minimum wage that eliminates monopsony power at one firm causes severe rationing at another. These results hold under an EITC and progressive labor income taxes calibrated to the U.S. economy.

JEL codes: E2, J2, J42

Keywords: Labor markets, Market structure, Oligopsony, Minimum wages

*Berger: Duke University. Herkenhoff: University of Minnesota. Mongey: Federal Reserve Bank of Minneapolis. This research was supported by the National Science Foundation (Award No. SES-2214431 and Award No. SES-2214460). Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily reflect the position of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

Minimum wage policies are widely implemented around the world, yet their utility is still the subject of debate. In *"The State of Labor Market Competition"* (2022), the U.S. Treasury identifies two main reasons to support a minimum wage: efficiency and redistribution.[1] The *efficiency* argument is that a minimum wage reduces monopsony power. The *redistribution* argument is that a minimum wage shifts resources towards lower income households. Quantifying each channel separately is important for understanding minimum wage policy.

In this paper, we extend our oligopsonistic model of labor markets (Berger, Herkenhoff, and Mongey, 2022, henceforth, BHM) and use it to conduct a quantitative analysis of the Federal minimum wage. The model is useful for such analysis as it captures redistributive motives as well as three key channels through which minimum wages can improve efficiency: (i) monopsony allows a higher minimum wage to raise wages and employment (*Direct effects*), (ii) oligopsony allows firms to respond to competitors paying a minimum wage (*Spillover effects*), and (iii) firm heterogeneity and granular markets allow reallocation from low to high productivity firms as the minimum wage binds (*Reallocation effects*). The model is quantitatively consistent with empirical evidence on these channels and hence a good laboratory for quantifying potential efficiency gains. We have two main results.

First, the efficiency gains from minimum wages are robustly small. We compute efficiency maximizing minimum wages using two methods. In a homogeneous worker environment redistributive motives are absent by construction and we compute an optimal minimum wage of $7.60. Gains are small: 0.2% of lifetime consumption and output increases 0.4%. These small gains that exist are equally attributable to competitors' responses via *Spillovers* and *Reallocation* of workers to more productive firms, while *Direct* effects are limited. Moreover, these gains are not small because there are no gains to be had. The potential welfare gains from eliminated monopsony power in the economy are large (6.3% of lifetime consumption), but a minimum wage is a poor tool for addressing inefficiency in labor markets. We repeat this exercise in an environment with worker heterogeneity where

[1] *"Raising the minimum wage is a straightforward approach to addressing lower wages under monopsony and can help increase employment."* (p.51, Efficiency), and then *"Raising the federal minimum wage would give nearly 32 million Americans a raise and would boost the purchasing power of low-income families ..."* (p.52, Redistribution)

redistributive motives are present. To abstract from redistribution we decompose welfare by combining elements of Floden (2001) and Dávila and Schaab (2022), and obtain an efficiency maximizing minimum wage is \$7.35. Efficiency gains remain small: 0.09% of lifetime consumption.

We find that efficiency gains are limited due to four forces that are germane in concentrated labor markets with heterogeneous firms: (1) The minimum wage bites most for near-competitive, low productivity firms who have little share of national employment[2]; (2) the range of employment-increasing minimum wages at low productivity firms is small since labor supply is elastic; (3) employment gains quickly become large employment losses as firms shrink beyond competitive levels of employment due to elastic firm demand; and (4) large firms that account for the most distortions raise their wages little in response to smaller, low wage competitors paying the minimum wage.

Second, a minimum wage can improve welfare overall via redistribution, at the expense of efficiency losses.[3] The extended model with worker heterogeneity includes both redistributive and efficiency motives. Under a Utilitarian objective, (i) the optimal minimum wage is \$11, (ii) but the welfare gains are only one-tenth of the potential gains from eliminating monopsony power (i.e. 2.8% whereas perfect competition yields gains of more than 30%), and (iii) 102.5% of the resulting welfare gains are driven by redistribution while efficiency is reduced by -2.5%.[4]

We find that redistribution via an EITC and progressive taxes consistent with U.S. policy does not negate these small welfare effects. Regarding efficiency, an EITC and progressive taxes exacerbate labor market power. This widens markdowns, which beckons a small increase in the optimal minimum wage. Regarding redistribution, a minimum wage redistributes from business owners to workers. Profits are largely unchanged under an EITC and progressive taxes, hence the re-

[2]The within-market, cross-sectional relationship between larger market shares and wider markdowns has been documented in the U.S. (Yeh, Macaluso, and Hershbein, 2022) and Denmark (Chan, Mattana, Salgado, and Xu, 2023). As described in these papers, our model is consistent with their facts.

[3]The potency of minimum wages to redistribute has been well documented (Derenoncourt and Montialoux, 2021; Cengiz, Dube, Lindner, and Zipperer, 2019). We show that our model generates spillovers up the wage distribution consistent with empirical evidence.

[4]The section on minimum wages in the Cahuc and Zylberberg (2004) textbook ends by questioning whether a minimum wage primarily acts through efficiency or redistribution. Our answer is: more than 100% through redistribution.

distributive role remains the same. Overall the optimum increases slightly, with similar gains.

We believe our model does not understate the channels through which minimum wages can generate efficiency gains. One reason is that our model replicates empirical evidence from the minimum wage literature: (i) *Direct effects*: Jardim et. al. (2022) and Azar, Huet-Vaughn, Marinescu, Taska, and von Wachter (2023); (ii) *Spillover effects*: Engbom and Moser (2022), (iii) *Reallocation effects*: Dustmann, Lindner, Schönberg, Umkehrer, and vom Berge (2022). Another reason is that small efficiency gains from a minimum wage hold across robustness exercises: (i) alternative labor supply elasticities, (ii) state-specific minimum wages in low and high income states, (iii) fixed capital and firm exit,[5] (iv) labor-labor substitution in production consistent with Katz and Murphy (1992) and Acemoglu and Autor (2011).

Our model necessarily omits a number of features that come to mind when thinking about the effects of minimum wages: pass-through to prices, automation, a non-unitary elasticity of substitution between capital and labor, inefficient rationing, and unemployment with incomplete markets. We conclude with a discussion of how each would likely lead to even smaller efficiency gains.

Literature. We analyze price controls in concentrated markets with strategic interactions between heterogeneous firms. Price controls in concentrated markets with strategic interaction between homogeneous firms has been studied in stylized cases (Molho, 1995; Reynolds and Rietzke, 2018; Bhaskar and To, 1999). Others study capacity constraints and rationing in competitive environments (de Palma, Picard, and Waddell, 2007; Ching, Hayashi, and Wang, 2015). We handle firm heterogeneity by expressing equilibrium conditions in terms of *shadow wages* which are *shadow markdowns* relative to marginal products. At the firm level, shadow markdowns encode (i) welfare losses from marginally tighter rationing under a minimum wage, and (ii) deviations from efficiency due to market power. We extend tools from BHM to aggregate these to an economy-wide shadow markdown, which narrows as a minimum wage erodes monopsonists' ability to set low wages, and then widens as employment is progressively rationed.

Recent, complementary, papers construct general equilibrium models with a

[5] This is a simplified version of exercises in putty-clay models of Aaronson, French, Sorkin, and To (2018) and Sorkin (2015).

minimum wage. Hurst et. al. (2022) study a search environment and putty-clay capital. They focus on positive outcomes and redistribution, with an expanded role for worker heterogeneity. We focus on normative outcomes and efficiency, with an expanded role for firm heterogeneity, which is necessary for incorporating empirically documented efficiency channels. Ahlfeldt et. al. (2022) computes welfare maximizing minimum wages in a spatial model of the German economy. Vogel (2022) finds that adding monopsony and a minimum wage to Katz and Murphy (1992) helps explain the evolution of the college wage premium. Haanwickel (2023) studies the effects of minimum wages on sorting and task assignment.

We study a neoclassical labor market, similar to Cahuc and Laroque (2014), Lee and Saez (2012) among others, while minimum wages have often been studied in frictional settings. Flinn (2006, 2010) documents the forces that shape optimal minimum wages in a frictional setting. Flinn and Mullins (2021) find that higher minimum wages lead firms to prefer renegotiation to wage-posting. Engbom and Moser (2022) extends Burdett and Mortensen (1998) to quantify the link between minimum wages and wage inequality, but do not consider what is optimal.

Overview. Section 1 extends BHM to include a minimum wage. Section 2 characterizes equilibrium behavior. Section 3 quantifies the efficiency maximizing minimum wage and small associated welfare gains. Section 4 adds worker heterogeneity. Section 5 quantifies the welfare maximizing minimum wage from a Utilitarian perspective. Section 6 repeats this exercise in the presence of taxes and transfers. Section 7 contains empirical replications, robustness and discussion of missing features. Section 8 concludes.

Additional proofs, derivations, figures and tables are contained in two appendices: (i) a *Supplemental (Online) Appendix*, published by this journal, and (ii) *Additional Materials to Minimum Wages Efficiency and Welfare* published as a separate working paper found on the authors' websites (Berger, Herkenhoff, and Mongey, 2024). We refer to these as Appendix O and Appendix A, respectively.

1 Homogeneous worker economy

Welfare gains from minimum wages in a homogeneous worker economy are an important benchmark as, by definition, they abstract from redistribution. We carefully describe our environment and equilibrium, since analysis of a minimum wage in a general equilibrium setting with firm heterogeneity is new. Section 3 provides our quantitative results.

Agents. Time is infinite and discrete, indexed by t. The economy consists of a single household and a continuum of firms. Firms are divided into a continuum of labor markets, $j \in [0,1]$. Each market has a fixed, finite number of firms M_j, $i \in \{1,\ldots,M\}$. Indices (i,j) identify a firm. Firms permanently differ in total factor productivity, z_{ij}. There is no entry. We later consider firm exit.

Goods and technology. Each firm produces a homogeneous good which trades in a perfectly competitive market at price P, normalized to one. Goods are used for consumption and investment. A firm rents capital k_{ij} and labor n_{ij} to produce output y_{ij} according to:

$$y_{ijt} = \overline{Z} z_{ij} \left(n_{ijt}^{\gamma} k_{ijt}^{1-\gamma} \right)^{\alpha} \quad , \quad \gamma \in (0,1] \quad , \quad \alpha > 0,$$

where $\overline{Z}$ is an aggregate productivity shifter. The production function has a unit elasticity of substitution between capital and labor.[6] We do not make a restriction that $\alpha < 1$, however this will be the case from the calibration of the model.

Labor market competition. With a finite number of firms in each local labor market, firms behave strategically. We assume Cournot competition: firms take as given the quantities of labor chosen by local competitors when taking their actions. Since labor market j is infinitesimal with respect to other labor markets, firms take quantities and wages outside of their labor market as given. We refer to this as Cournot oligopsony. Because firms are oligopsonists, they earn profits, $\pi_{ij} \geq 0$. Total profits, Π, are rebated to the household.

Minimum wages and rationing constraints. Denote the minimum wage $\underline{w} \geq 0$. Like any neoclassical economy with price controls, for certain levels of the minimum wage, there may be excess supply of labor to a firm: at $\underline{w}$ workers want to supply more labor than a firm demands. Since the labor market for a given firm may not necessarily clear for a given minimum wage, we allow each firm to specify a constraint $\overline{n}_{ij}$. This is a sign on the firm's door telling the household the

[6] There are range of estimates of the elasticity of substitution between capital and labor reported in the empirical literature, however, most estimate elasticities in the range of 0.7 to 1.2. See Section 7 for further discussion to our baseline assumption.

maximum amount of labor the firm is willing to hire, hence $n_{ij} \leq \overline{n}_{ij}$. We call this a *rationing constraint*.

1.1 Household problem

Given initial capital K_t, the household chooses next period capital K_{t+1} and the allocation of labor $\{n_{ijt}\}$ across firms. It takes as given the rationing constraints $\{\overline{n}_{ijt}\}$, wages $\{w_{ijt}\}$, the rental rate of capital R_t, and profits Π_t. Households have concave preferences over consumption and a convex disutility of labor. Labor disutility has a nested-*CES* functional form, taken directly from BHM and discussed in detail below. Since the household's problem is dynamic, we add time subscripts to the variables in this section.

Household preferences are given by,

$$\mathcal{U} = \sum_{t=0}^{\infty} \beta^t u\Big(C_t, N_t\Big) = \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{1}{\overline{\varphi}^{1/\varphi}} \frac{N_t^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}\right], \tag{1}$$

$$\text{where} \quad C_t := \int_0^1 \sum_{i=1}^{M_j} c_{ijt}\, dj \quad , \quad N_t := \left[\int_0^1 n_{jt}^{\frac{\theta+1}{\theta}}\, dj\right]^{\frac{\theta}{\theta+1}} \;,\; n_{jt} := \left[\sum_{i=1}^{M_j} n_{ijt}^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}}.$$

As in BHM, we assume elasticities of substitution η and θ are such that the household finds jobs within a market to be closer substitutes than across markets, i.e. $\eta > \theta$. This means labor supply to firms is more elastic with respect to within-market wage differences across firms, relative to across-market wage differences. The intuition is that η captures intra-market frictions (e.g. commute costs) where as θ captures inter-market frictions (e.g. moving costs). As $\eta \to \infty$, intra-market frictions approach zero, and firms within a market are perfect substitutes: the household only sends workers to the firm that offers the highest wage. As $\theta \to \infty$, inter-market frictions approach zero, and markets are perfect substitutes: the household only sends workers to the market that offers the highest wage. Neoclassical monopsony is nested under $\eta = \theta$, which we later exploit to isolate mechanisms. Finally, note that household labor supply features wealth effects. Empirically, wealth effects are important for labor supply in the U.S. (Golosov, Graber, Mogstad, and Novgorodsky, 2021). Quantitatively, including these are important as the minimum wage will effect, labor, capital and profit income. The parameter $\overline{\varphi}$, along with the shifter $\overline{Z}$ in the production function, provide normalizing constants that we will calibrate to match properties of the levels of employment and wages in the economy.

In addition to labor income, the household earns capital income and profits, and chooses how much to consume and invest. Their budget constraint is:

$$C_t + K_{t+1} = \int \sum_{i=1}^{M_j} w_{ijt} n_{ijt} \, dj + R_t K_t + (1-\delta) K_t + \Pi_t. \tag{2}$$

Given prices, the household's problem is to choose labor n_{ijt} and capital K_{t+1} to maximize utility (10) subject to (11) and labor rationing constraints, $n_{ijt} \leq \overline{n}_{ijt}$.

Household labor supply curve. Let $\beta^t \nu_t$ be the multiplier on the household's budget constraint. We write the multiplier on the rationing constraint as $\zeta_{ijt} = \beta^t \nu_t w_{ijt} \left(1 - p_{ijt}\right)$. This way, the first order condition for labor supply equates the usual product of marginal rates of substitution to $w_{ijt} p_{ijt}$:

$$w_{ijt} p_{ijt} = \underbrace{\left(\frac{n_{ijt}}{n_{jt}}\right)^{\frac{1}{\eta}}}_{MRS \text{ b/w firms}} \underbrace{\left(\frac{n_{jt}}{N_t}\right)^{\frac{1}{\theta}}}_{MRS \text{ b/w markets}} \underbrace{\left(\frac{-u_n\left(C_t, N_t\right)}{u_c\left(C_t, N_t\right)}\right)}_{MRS \text{ b/w } C \text{ and } N} \quad , \quad \underbrace{\zeta_{ijt}\left(\overline{n}_{ijt} - n_{ijt}\right) = 0}_{\text{Complementary slackness}} \tag{3}$$

The normalized multiplier $p_{ijt} \in (0,1]$, and $p_{ijt} < 1$ if and only if the rationing constraint binds, giving the wedge between the price paid for labor and the household's marginal rate(s) of substitution.[7]

We can combine conditions (3) to obtain an *inverse labor supply schedule*:

$$w\left(n_{ijt}, \overline{n}_{ijt}, n_{jt}, \mathbf{S}_t\right) = \begin{cases} \left(\frac{n_{ijt}}{n_{jt}}\right)^{\frac{1}{\eta}} \left(\frac{n_{jt}}{N_t}\right)^{\frac{1}{\theta}} \left(\frac{-u_n(C_t,N_t)}{u_c(C_t,N_t)}\right) & , n_{ijt} \in \left[0, \overline{n}_{ijt}\right) \\ \in \left[\left(\frac{\overline{n}_{ijt}}{n_{jt}}\right)^{\frac{1}{\eta}} \left(\frac{n_{jt}}{N_t}\right)^{\frac{1}{\theta}} \left(\frac{-u_n(C_t,N_t)}{u_c(C_t,N_t)}\right), \infty \right) & , n_{ijt} = \overline{n}_{ijt} \end{cases} \tag{4}$$

Taking as given aggregates $\mathbf{S}_t$, and competitors' employment which enters n_{jt}, when a firm chooses n_{ijt} and $\overline{n}_{ijt}$, (4) gives the wage that will have to be paid. Appendix O.D provides additional details on the derivation, and shows that at $\overline{n}_{ijt}$, the households' labor supply schedule is a correspondence. A firm would never pay more than the minimum wage necessary to deliver $\overline{n}_{ijt}$ workers, allowing us to work with a one-to-one function over $n_{ijt} \in [0, \overline{n}_{ijt}]$.

Household investment. The household's Euler equation implies that steady-state household capital supply that is perfectly elastic at $R = 1/\beta + (1-\delta)$.

[7] Throughout we use *binding* to mean a strictly binding constraint ($\zeta_{ijt} > 0$, $n_{ijt} = \overline{n}_{ijt}$), and *slack* to indicate a weakly slack constraint ($\zeta_{ijt} = 0$, $n_{ijt} \leq \overline{n}_{ijt}$).

1.2 Firm problem

Firm i in market j takes as given local competitors' employment n_{-ijt} as well as aggregates $\mathbf{S}_t$ and chooses its (i) wage w_{ijt}, (ii) employment n_{ijt}, (iii) capital k_{ijt}, and (iv) rationing constraint $\overline{n}_{ijt}$ in order to maximize profits.

The firm faces several constraints. They must respect the minimum wage $w_{ijt} \geq \underline{w}$, their self-imposed rationing constraint $n_{ijt} \leq \overline{n}_{ijt}$ as well as the household's inverse labor supply schedule $w_{ijt} = w\left(n_{ijt}, \overline{n}_{ijt}, n_{jt}, \mathbf{S}_t\right)$ which depends on local competitors' employment through n_{jt}.

Therefore the firm problem is given by,

$$\max_{\overline{n}_{ijt}, n_{ijt}, w_{ijt}, k_{ijt}} \overline{Z} z_{ijt} \left(n_{ijt}^{\gamma} k_{ijt}^{1-\gamma}\right)^{\alpha} - R_t k_{ijt} - w_{ijt} n_{ijt} \tag{5}$$

$$\text{subject to} \qquad w_{ijt} \geq \underline{w} \ , \ n_{ijt} \leq \overline{n}_{ijt} \ , \ w_{ijt} = w\Big(n_{ijt}, \overline{n}_{ijt}, n_{jt}(n_{ijt}, n_{-ijt}), \mathbf{S}_t\Big).$$

Under Cournot competition, the firm understands $\partial w\left(n_{ijt}, \overline{n}_{ijt}, n_{jt}, \mathbf{S}_t\right) / \partial n_{ijt} \neq 0$ and that $\partial n_{jt} / \partial n_{ijt} \neq 0$, yielding *oligopsonistic* wage setting. In particular, the firm understands that their hiring affects the wage they pay (i) directly and (ii) indirectly through market level employment n_{jt}:

$$n_{jt}\Big(n_{ijt}, n_{-ijt}\Big) := \left[n_{ijt}^{\frac{\eta+1}{\eta}} + \sum_{k \neq i}^{M_j} n_{kjt}^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}} \qquad , \qquad \left.\frac{\partial n_{jt}\left(n_{ijt}, n_{-ijt}\right)}{\partial n_{ijt}}\right|_{n_{-ijt}} \neq 0.$$

For ease of exposition in subsequent sections, we first optimize out firm capital. The resulting firm profit function is given by $\pi_{ijt} = \widetilde{Z} \widetilde{z}_{ijt} n_{ijt}^{\widetilde{\alpha}} - w_{ijt} n_{ijt}$, where

$$\widetilde{Z} := \overline{Z}^{\frac{1}{1-(1-\gamma)\alpha}} \ , \ \widetilde{\alpha} := \frac{\gamma\alpha}{1-(1-\gamma)\alpha} \ , \ \widetilde{z}_{ijt} := \Big[1-(1-\gamma)\,\alpha\Big] \left(\frac{(1-\gamma)\,\alpha}{R_t}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}} z_{ijt}^{\frac{1}{1-(1-\gamma)\alpha}}.$$

1.3 Equilibrium

We focus on a steady-state equilibrium. An *oligopsonistic Nash-Cournot* steady-state equilibrium consists of prices, aggregates (profits, market and national employment indices), household and firm policy functions such that: (1) given prices and aggregates, household policy functions characterizing labor supply and capital supply are optimal, (ii) given national aggregates, market competitors' employment and household labor supply functions, firm employment, capital, and rationing decisions are optimal, (iii) labor and capital markets clear.

2 Characterization of firm and market behavior

In this section we describe how minimum wages constrain firms' wage setting, show how a formulation of optimality conditions in terms of *shadow wages* can be used to gain tractability and aggregate, describe firm's optimal response to a minimum wage in partial equilibrium, and then how firms' equilibrium responses to competitors shape the equilibrium of a particular labor market. This produces the *Direct*, *Spillover* and *Reallocation* channels, discussed in the Introduction, through which a minimum wage may prove efficiency in a concentrated labor market. We proceed via illustrative and numerical examples drawn from the model as calibrated in the following Section.

2.1 Preliminaries

We start with some preliminaries. Proofs for all statements in this Section may be found in Appendix O.D. Since the firm's problem is static, we omit time subscripts. We begin by defining three regions of the firm's problem, for which we will derive optimality conditions. Under successively higher minimum wages, a firm moves through these regions

- *Region I*: Firm is unconstrained by $\underline{w}$, household is on its labor supply curve.
- *Region II*: Firm is constrained by $\underline{w}$, household is on its labor supply curve.
- *Region III*: Firm is constrained by $\underline{w}$, household is off its labor supply curve.

Firm wage setting with a zero minimum wage. When $\underline{w} = 0$, the firm problem is identical to BHM. Rationing constraints are irrelevant and wages are a variable markdown μ_{ij} on the marginal revenue product of labor,

$$w_{ij} = \mu_{ij}\widetilde{\alpha}\widetilde{z}_{ij}n_{ij}^{\widetilde{\alpha}-1} \quad , \quad \mu_{ij} = \frac{\varepsilon_{ij}}{\varepsilon_{ij}+1} \quad , \quad \varepsilon_{ij} = \left[\frac{1}{\eta} + \left(\frac{1}{\theta} - \frac{1}{\eta}\right) s_{ij}\right]^{-1} \quad , \quad s_{ij} = \frac{w_{ij}n_{ij}}{\sum_i w_{ij}n_{ij}}. \tag{6}$$

Here, ε_{ij} is the perceived labor supply elasticity of firm ij which depends on the firm's wage-bill share s_{ij}. If a firm is by itself in a market, $s_{ij} = 1$, and its perceived labor supply elasticity is θ. Intuitively, a solo monopsonist making a marginal hire understands it must draw workers from outside its market. If a firm is atomistic, $s_{ij} = 0$, its perceived labor supply elasticity is η. To a tiny firm, local and national labor markets are equally massive, and hence the relevant elasticity is intra-market.

The market equilibrium in BHM is a simple fixed point in wage-bill shares s_{ij}. This is not the case when $\underline{w} > 0$.

Firm wage setting with a minimum wage. When $\underline{w} > 0$ some firms' wages are not optimal (Region II), while others' wages are not allocative (Region III). Equations (6) do not hold, which makes analysis and aggregation intractable. Hence, we next develop a representation of our economy that mimics (6) but in terms of allocative *shadow wages* and *shadow markdowns*. This accommodates aggregation and decomposition of the optimal minimum wage.[8]

2.2 Characterization using shadow wages

We show that recasting the equilibrium conditions for firms' optimal wages and employment in terms of *shadow wages* allow us to (i) succinctly analyze firm behavior, and (ii) aggregate optimality conditions in the absence of market clearing to study general equilibrium, which (iii) allows us to pinpoint efficiency gains and losses due to minimum wages. Using our normalized multiplier p_{ij}, we define a shadow wage that admits aggregation.

***Definition*:** The *shadow wage, markdown and wage-bill share* $\{\widetilde{w}_{ij}, \widetilde{\mu}_{ij}, \widetilde{s}_{ij}\}$ are:

$$\widetilde{w}_{ij} := p_{ij}w_{ij} = \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} \left(\frac{-u_n\left(C_t, N_t\right)}{u_c\left(C_t, N_t\right)}\right) \quad , \quad \widetilde{\mu}_{ij} := \frac{\widetilde{w}_{ij}}{\widetilde{\alpha}\widetilde{z}_{ij}n_{ij}^{\widetilde{\alpha}-1}} \quad , \quad \widetilde{s}_{ij} := \frac{\widetilde{w}_{ij}n_{ij}}{\sum_{i=1}^{M_j}\widetilde{w}_{ij}n_{ij}}.$$

The *shadow wage* captures two ideas. First, it is the relevant allocative price for household employment in that it always places the household on its supply curve. Second, since $\widetilde{w}_{ij} = p_{ij}w_{ij} \leq w_{ij}$, then $\widetilde{w}_{ij}$ encodes the bindingness of the rationing constraint. The *shadow markdown* is the ratio of the shadow wage to the worker's marginal revenue product of labor. Since shadow wages determine quantities, and firms care about competitors' quantities, the relevant market share for a firm is its *shadow share*. This is higher ($\widetilde{s}_{ij} > s_{ij}$) when competitors' shadow wages are lower than their actual wages ($\widetilde{w}_{ik} < w_{ik}$).

Using these definitions we rewrite the firm's optimal wage and employment decisions in terms of shadow wages in Regions I, II, and III.

[8] This approach has been adopted in extensions of this paper to include migration (Marhsall, 2023) and firm organizational structure (Janez and Delgado-Prieto, 2023).

Region I: For firms in Region I, $\underline{w}$ is not binding, so the rationing constraint is not binding: $p_{ij} = 1$, $\widetilde{w}_{ij} = w_{ij}$ and $\widetilde{\mu}_{ij} = \mu_{ij}$. However, the firm's markdown and wage in equation (6) are now written in terms of the shadow wage-bill share (proof in Appendix O.D):

$$\widetilde{w}_{ij} = \widetilde{\mu}_{ij}\widetilde{\alpha}\widetilde{z}_{ij}n_{ij}^{\widetilde{\alpha}-1} \ , \ \widetilde{\mu}_{ij} = \frac{\varepsilon_{ij}}{\varepsilon_{ij}+1} \ , \ \varepsilon_{ij} = \left[\frac{1}{\eta} + \left(\frac{1}{\theta} - \frac{1}{\eta}\right)\widetilde{s}_{ij}\right]^{-1} \ , \ \widetilde{s}_{ij} = \left.\frac{\partial \log n_j(n_{ij}, n_{-ij})}{\partial \log n_{ij}}\right|_{n_{-ij}} . \quad (7)$$

In Region I, employment n_{ij} can be read off of the household's labor supply curve. The novelty is its expression in terms of shadow wages and shadow wage indices at the market and aggregate level. Hence our formulation admits aggregation:

$$n_{ij} = \left(\frac{\widetilde{w}_{ij}}{\widetilde{w}_j}\right)^{\eta}\left(\frac{\widetilde{w}_j}{\widetilde{W}}\right)^{\theta} N \ , \ \widetilde{w}_j := \left[\sum_{i\in j}\widetilde{w}_{ij}^{1+\eta}\right]^{\frac{1}{1+\eta}} \ , \ \widetilde{W} := \left[\int \widetilde{w}_j^{1+\theta} dj\right]^{\frac{1}{1+\theta}} \ , \ N = \overline{\varphi}\widetilde{W}^{\varphi}C^{-\sigma\varphi} \quad (8)$$

The key tractability issue of working with the minimum wage is that it is not an allocative price (for example, in a particular two firms could have the same employment while one is unconstrained and another is paying a minimum wage). Equations (8) show that the shadow wage is allocative, uniquely determining firm employment. This then remains true as we aggregate to the market and economy level. In fact, the aggregate supply curve is instantly recognizable as labor supply under MaCurdy (1981) preferences with wealth effects, but with the aggregate shadow wage $\widetilde{W}$ taking the role of the allocative price of labor. This encodes the full distribution of multipliers across all firms. Solving the model requires having a notion of prices at the market and aggregate level, and hence the shadow wage representation facilitates solving the model.

Region II: The firm is constrained by the minimum wage but the household is on their labor supply curve and so the rationing constraint is not binding: $p_{ij} = 1$, $\widetilde{w}_{ij} = \underline{w}$, $\widetilde{\mu}_{ij} = \frac{\underline{w}}{\widetilde{\alpha}\widetilde{z}_{ij}n_{ij}^{\widetilde{\alpha}-1}}$. Employment n_{ij} is given by the household's labor supply curve in equation (8) evaluated at $\widetilde{w}_{ij} = \underline{w}$. As the minimum wage increases, $\widetilde{\mu}_{ij}$ increases (i.e. markdowns narrow). At the border of Regions II and III, the wage and marginal revenue product are equalized, hence—at the firm level—the employment allocation is efficient.

Region III: The firm is constrained by the minimum wage, the household is off their labor supply curve and the rationing constraint binds: $p_{ij} < 1$, $w_{ij} = \underline{w} = mrpl_{ij} = \widetilde{\alpha}\widetilde{z}_{ij}n_{ij}^{\widetilde{\alpha}-1}$, and hence $\widetilde{\mu}_{ij} = p_{ij}$. As the minimum wage increases the rationing constraint binds further, and the associated inefficiency is encoded in a wider shadow markdown.

Finally, to solve for the optimal rationing constraint, note that a firm would never hire at a point where their marginal revenue product is below the minimum wage (proof in Appendix O.D). Intersecting $mrpl_{ij}$ and $\underline{w}$ gives:

$$\overline{n}_{ij} = \overline{n}\left(\widetilde{z}_{ij}, \underline{w}\right) = \left(\frac{\widetilde{\alpha}\widetilde{z}_{ij}}{\underline{w}}\right)^{\frac{1}{1-\widetilde{\alpha}}} .$$

In Region III this is optimal, and weakly optimal in Regions I and II, where the constraint is slack. Importantly, in Region III, $n_{ij} = \overline{n}_{ij}$ implies the household does not send surplus labor to firm-ij. There is no idle *excess supply of labor* as in the neoclassical presentation of the minimum wage. Workers that would work at firm-ij at $\underline{w}$—if they were demanded—observe $\overline{n}_{ij}$ and go work elsewhere. The rationing constraint is naturally independent of local competitors' employment levels, which maintains tractability.

2.3 Firm response to minimum wage - Partial equilibrium

To clarify the above, Figure 1 illustrates firms' optimality conditions in partial equilibrium in a single market j (i.e. holding all other firms' wages and employment fixed). To reduce clutter, we omit the market subscript j.

Panel A reproduces the firm's optimality condition in a neoclassical monopsony model without a minimum wage.[9] With monopsony power, employment n_i^0 is below the competitive benchmark n_i^c, with lower wages $w_i^0 < w_i^c$.

In Panel B, a non-binding minimum wage is introduced. The firm takes as given the inverse labor supply schedule (4), which emerges from household optimality and maps choices of $(n_i, \overline{n}_i)$ into w_i. The firm's optimal rationing constraint $\overline{n}_i = \overline{n}(\underline{w}, \widetilde{z}_i)$, (equation 2.2) truncates labor supply, and is slack. The firm's optimal employment is unaffected by $\underline{w}$ and the shadow wage and shadow markdown coincide with Panel A.

In Panel C, a higher minimum wage pushes the firm into Region II: the minimum wage now binds, and optimal employment is pinned down by household

[9]If the downward sloping marginal revenue product of labor reflected diminishing marginal revenue—as would be the case for a monopolistically competitive producer—the second component of profits would be due to a price markup.

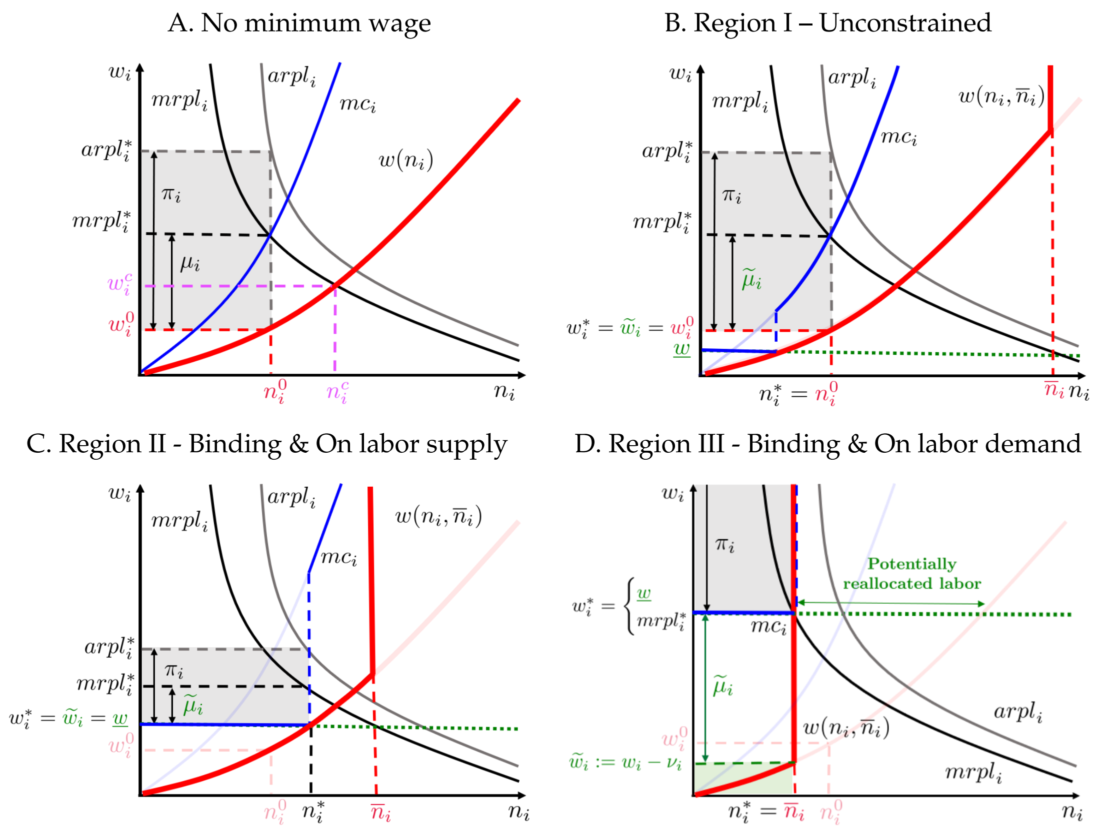

Figure 1: Increase in $\underline{w}$ - Partial equilibrium

Notes: The long dotted horizontal line corresponds to the minimum wage $\underline{w}$. The red line (marked $w(n_i, \overline{n}_i)$), gives the household's *inverse labor supply schedule* $w(n_i, \overline{n}_i, N)$, which depends on its labor supply and the rationing constraint $\overline{n}_i$. The blue line (marked mc_i) gives the firm's *marginal cost of labor* along its *perceived labor supply curve* $\max\{\underline{w}, w(n_i, \overline{n}_i, N)\}$ on $n_i \in (0, \overline{n}_i]$.

labor supply. Relative to Panel B, wages and employment are higher, and the loss in profits is born by the firm.[10] The optimal rationing constraint remains slack ($p_i = 1$), and the shadow and minimum wage coincide. Increasing $\underline{w}$ would further narrow the firm's shadow markdown $\widetilde{\mu}_i$. This represents the *Direct channel* through which a higher minimum wage can improve efficiency by narrowing $\widetilde{\mu}_i$.

Increasing $\underline{w}$ further pushes the firm past the efficient allocation ($\widetilde{\mu}_i = 1$) and into *Region III* (Panel D). At $(\underline{w}, \overline{n}_i)$, the marginal disutility of labor—read off the supply curve—is below the wage. The shadow markdown $\widetilde{\mu}_i$ measures this inefficiency. Note that $\overline{n}_i$ is less than the initial n_i^0: the minimum wage has lead to less

[10] In Region II, the marginal cost curve is different from the benchmark economy. The new marginal cost curve is horizontal and equal to $\underline{w}$ until it reaches the labor supply curve. Up to this point workers are paid $\underline{w}$. Marginal cost then jumps. Above the minimum wage, hiring an additional worker requires increasing pay for all existing workers. As marginal cost jumps above marginal revenue, profit maximizing employment is on the labor supply curve at $\underline{w}$.

efficient employment than a baseline with market power.

Under the 'textbook' treatment of the minimum wage, firms are homogeneous and one could label the gap at $\underline{w}$ between labor demand and supply as *non-employment* generated by the minimum wage. A novel feature of our economy is firm heterogeneity. Rationing constraints in this economy are important. Rather than having idle labor outside firm i, workers understand labor is rationed, and can be productively reallocated to other firms within- and across-markets. Since low productivity firms will be the first to enter Region III, and reallocation is more elastic within rather than across markets, this reallocation will primarily be to more productive firms in market j. This represents the *Reallocation channel* through which a higher minimum wage can improve efficiency: jobs aren't necessarily destroyed, they're partially reallocated.

At the microeconomic level of the firm, endogenous rationing constraints deliver a clear picture of the wages and shadow wages that rationalize equilibrium employment. Shadow markdowns capture inefficiencies due to (i) market power in Region I, (ii) diminished market power in Region II, and (iii) binding rationing constraints due to the minimum wage in Region III. We now show how these objects characterize the efficiency effects of the minimum wage at the market level.

2.4 Market response to minimum wage

We now consider the same comparative static but in a market equilibrium, this time holding aggregates outside of the market fixed. In simple monopsony models the only channel through which minimum wages improve efficiency is via the *Direct channel* of moving firms toward their competitive wage in Region II. The *market equilibrium* of our oligopsony model delivers two additional channels: *Spillovers* and *Reallocation*. In Section 7 we describe empirical evidence for these channels, and show how our model quantitatively reproduces this evidence. Figure 2 plots a numerical example of a market with three firms, using our calibrated model (for details see figure footnote).

Channel I - Direct. The red, dotted, line describes the low productivity firm's movement through the three regions described in Figure 1. Its wage increases one-for-one with $\underline{w}$ across Regions II and III. The *Direct* efficiency gain is shown

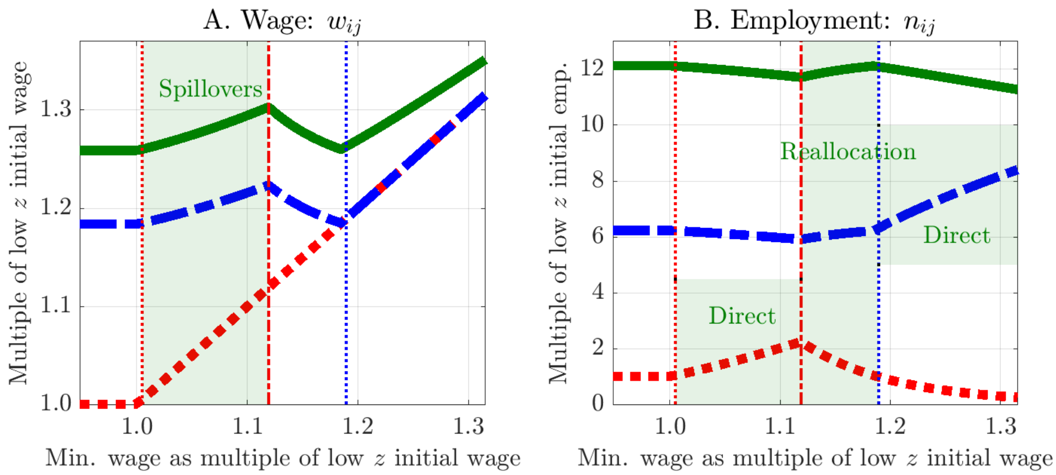

Figure 2: Increase in $\underline{w}$ - Firm outcomes in market equilibrium

Notes: All aggregates are held fixed and we plot outcomes for a market with three firms as the minimum wage is increased. The x-axis plots the minimum wage relative to unconstrained optimal wage of the low productivity firm: $\underline{w}/w^*_L$. We increase the minimum wage from 10 percent below to 50 percent above this wage. This figure is produced using parameters from 1. $M_j = 3$ and the productivities are given by $z^{low} = 1.97$ (dotted, red), $z^{med} = 4.04$ (dashed, blue), $z^{high} = 6.42$ (solid, green). National W and N are held fixed at value corresponding to $\underline{w} = 0$.

by the first shaded region in Panel B: employment increases in Region II. Wages and employment of the medium (blue) and high (green) productivity firms reflect the Nash equilibrium at the market level. These firms are larger, and pay higher wages. With large market shares, they face less elastic supply, so their wages are wider markdowns on their marginal product of labor (equation 6).

Channel II - Spillovers. As the low productivity firm's wage increases in Region II, its market share increases, which puts pressure on the shares of the unconstrained firms. Facing stiffer competition, the unconstrained firms' equilibrium markdowns narrow (equation 6). Their wages consequently increase in the shaded region in Panel A. This *Spillover* effect has positive implications for efficiency. While the minimum wage only binds for the low productivity firm, all firms' equilibrium markdowns are narrowing. The elasticity of firms' wages to competitors' is therefore a key determinant of the efficiency properties of minimum wages.

Channel III - Reallocation. As the minimum wage increases, the *Direct* gains at the low productivity firm are undone: its employment shrinks in Region III. However, the high elasticity of substitution of labor within- relative to across-markets implies that these employment losses are largely reallocated to its discretely more productive competitors. This is a third form of efficiency gain. In Section 5 we repeat this exercise under $\theta = \eta$. Reallocation is completely neutralized, as cuts by the low productivity firm spread out across all markets. The reallocation of employment from lower to higher productivity firms within markets is therefore also a key determinant of the efficiency properties of minimum wages.

2.5 Aggregation

To say something about overall efficiency we need to aggregate these effects. At the market level, output y_j, employment n_j and the market shadow wage $\widetilde{w}_j$ are jointly determined by (proof see Appendix A.G):

$$\underbrace{y_j = \omega_j \widetilde{z}_j n_j^{\widetilde{\alpha}}}_{\text{1. Output}} \quad , \quad \underbrace{\widetilde{w}_j = \widetilde{\mu}_j \times \widetilde{\alpha} \widetilde{z}_j n_j^{\widetilde{\alpha}-1}}_{\text{2. Shadow wage}} \quad , \quad \underbrace{\widetilde{n}_j = \left(\frac{\widetilde{w}_j}{\widetilde{W}}\right)^{\theta} N}_{\text{3. Labor supply}} .$$

The wedges $\widetilde{z}_j$, $\widetilde{\mu}_j$ and ω_j depend only on the joint distribution of $\{\widetilde{z}_{ij}, \widetilde{\mu}_{ij}\}_{j=1}^{M_j}$:

$$\underbrace{\widetilde{z}_j := \left[\sum_{i \in j} \widetilde{z}_{ij}^{\frac{1+\eta}{1+\eta(1-\widetilde{\alpha})}}\right]^{\frac{1+\eta(1-\widetilde{\alpha})}{1+\eta}}}_{\text{1. Market productivity}} \quad , \quad \underbrace{\widetilde{\mu}_j := \left[\sum_{i \in j} \left(\frac{\widetilde{z}_{ij}}{\widetilde{z}_j}\right)^{\frac{1+\eta}{1+\eta(1-\widetilde{\alpha})}} \widetilde{\mu}_{ij}^{\frac{1+\eta}{1+\eta(1-\widetilde{\alpha})}}\right]^{\frac{1+\eta(1-\widetilde{\alpha})}{1+\eta}}}_{\text{2. Market shadow markdown}} \quad , \quad \underbrace{\omega_j := \sum_{i \in j} \left(\frac{\widetilde{z}_{ij}}{\widetilde{z}_j}\right)^{\frac{1+\eta}{1+\eta(1-\widetilde{\alpha})}} \left(\frac{\widetilde{\mu}_{ij}}{\widetilde{\mu}_j}\right)^{\frac{\eta\widetilde{\alpha}}{1+\eta(1-\widetilde{\alpha})}}}_{\text{3. Market misallocation}} .$$

The shadow wage representation isolates the channels through which minimum wages affect efficiency. In the efficient allocation all markdowns are equal to one, implying $(\widetilde{\mu}_j, \omega_j) = (1, 1)$. Hence, the terms $(\widetilde{\mu}_j, \omega_j)$ encode deviations from the efficient allocation. Note that $\widetilde{\mu}_j$ exists with or without variable markdowns. It captures the neoclassical markdown distortions that are present in monopsonistic frameworks without firm heterogeneity (e.g. Robinson, 1933). The term ω_j only exists in environments with firm heterogeneity. It captures misallocation and encodes the interaction between firm heterogeneity, market power and minimum wages. It is smaller when more productive firms operate with wider (shadow) markdowns, which is the case in our oligopsony environment when the minimum wage is zero. When minimum wages are binding, shadow markdowns widen at low productivity firms pushed into Region III, which can potentially relieve some of the misallocation in the baseline economy.

Figure 3 shows how market aggregate wedges $(\widetilde{\mu}_j, \omega_j)$ evolve in the numerical

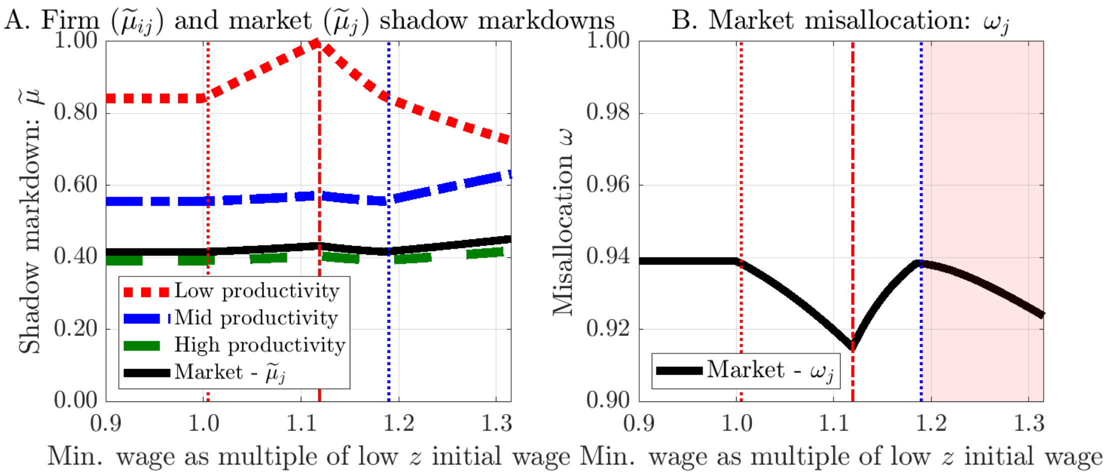

Figure 3: Increase in $\underline{w}$ - Market outcomes - Shadow markdown and misallocation

Notes: The model economy is identical to Figure 2. The productivities are given by $z^{low} = 1.97$ (dotted, red), $z^{med} = 4.04$ (dash-dot, blue), $z^{high} = 6.42$ (long dash, green). Panel *A* plots the market markdown $\widetilde{\mu}_j$ (solid, black). Panel *B* plots the market misallocation $\widetilde{\omega}_j$ (solid, black). Moving from left to right, the first vertical dotted line corresponds to the low productivity firm moving from Region I to II (red dotted), the next corresponds to the move from Region II to III (red dash-dot), and the third line corresponds to the medium productivity firms moving from Region I to II (blue dotted).

example from Figure 2. We note two results. First, productivity weighting in $\widetilde{\mu}_j$ implies that the market shadow-markdown is shaped by the *Spillover* responses of unconstrained firms (Panel A), rather than the *Direct* effect via the narrowing of the low producitivity firm's markdown. The model has a potentially strong role for spillovers in shaping efficiency. Second, misallocation has ambiguous effects (Panel B). Indeed, misallocation improves while the low productivity firm (red, dotted) is in Region III and its competitors are unconstrained. However, it worsens once the medium productivity firm starts paying the minimum wage (shaded). The high productivity firm (long dash, green) responds by increasing its wage less than one-for-one, so employment is reallocated down the productivity ladder, worsening ω_j, lowering output.

Taking stock. A key take-away from Figures 2 and 3 is that empirical evidence of any channel may not extend more generally. First, *Direct* gains only occur in the window of Region II, and are down-weighted as they are mostly incurred at low productivity firms. Second, *Spillovers* are moderated by large firms responding little to small firms' wage increases. Third, *Reallocation* cuts both ways as Region II growth comes at the expense of employment at more productive firms. Firm heterogeneity and strategic interactions provide the mechanics through which each channel operates. Yet when aggregated, efficiency gains and losses may offset. These rich interactions necessitate a quantitative general equilibrium approach that aggregates across many markets that are distributed across the spectrum of these effects. The remainder of our analysis seeks to implement this.

3 Homogeneous worker results

We calibrate our homogeneous worker economy and compute the efficiency gains from minimum wages. The key benefit of this environment is that it isolates efficiency since, by definition, there is no redistribution. We find efficiency gains from minimum wages are small and limited by firm heterogeneity. This headline result will be robust to adding rich household heterogeneity (Section 4).

3.1 Calibration

We calibrate the economy to US data, using a combination of Census data, Bureau of Labor Statistics (BLS), and Current Population Survey (CPS). In particular, our calibration uses moments based on the Longitudinal Business Database (LBD) released by our prior work (BHM). LBD data is from 2014, the latest data available to BHM. We use pre-Covid 2019 data from the CPS. Parameters and moments are summarized in Table 1.

We externally calibrate parameters in Table 1A. Discounting implies a risk free rate of 4 percent annually (β). Depreciation is 10 percent (δ). Curvature in marginal utility of consumption is 1.05 (σ), so approximately log, and the Frisch elasticity of aggregate labor supply is 0.62 (φ).[11]

The distribution of firms across markets matches LBD data. Markets are treated as in BHM as a combination of a NAICS 3-digit industry and a commuting zone. A firm in the data is a collection of all establishments with the same *firmid* in the commuting zone. We compute total employment and average worker wages across these establishments. The distribution of firms across markets $M_j \sim G(M)$ is comprised of a mass point of 0.09 at $M_j = 1$ and a generalized Pareto distribution for

[11] Given σ we use recent evidence to infer φ by combining (i) estimates on marginal propensities to consume and earn from Golosov et. al. (2022), and (ii) data on the average propensity to consume from the BLS. Details are in Appendix A.E.

Parameters		Value	Moment and source	Value
A. External				
Risk free rate	r	0.04		
Depreciation rate	δ	0.10		
Coefficient of risk aversion	σ	1.05		
Aggregate Frisch elasticity	φ	0.62		
Number of markets	J	5,000		
Distribution of number of firms	$G(M_j)$		Pareto with mass point at $M_j = 1$	
			Mean, variance, skewness of distribution	
			15 percent of markets have 1 firm	
Across market substitutability	θ	0.42	Estimate from BHM (2021)	
Within market substitutability	η	10.85	Estimate from BHM (2021)	
B. Internally estimated				
Productivity dispersion Std$[\log z_{ij}]$	σ_z	0.312	Payroll weighted $\mathbb{E}[HHI^{wn}]$ (LBD)	0.11
Decreasing returns in production	α	0.940	Labor share	0.57
Labor exponent in production	γ	0.808	Capital share	0.18
Labor disutility shifter	$\overline{\varphi}$	9.11×10^{11}	Average firm size	22.8
Productivity shifter	$\widetilde{Z}$	11.73	Binding at \$15 (CPS, %)	30.6

Table 1: Calibration of common parameters

$M_j > 1$. Tail, shape and location parameters are chosen to best match the mean (113.1), standard deviation (619.0) and skewness (26.1) of the empirical distribution of M_j in the LBD. We solve the model with $J = 5,000$ markets.

Preference parameters (θ, η) are taken from BHM. With $M_j < \infty$, firms exercise market power in their local labor markets. If $\eta > \theta$, labor supply is more elastic within- than across- markets, and firms with a larger market share will be less responsive to shocks. BHM uses the relative response of firms with large and small market shares following shocks to the marginal revenue product of labor to identify θ and η: $(\theta, \eta) = (0.42, 10.85)$. Below we show that under $\theta = \eta = 3.02$—which delivers the same labor share as the baseline economy but without oligopsony—efficiency gains from minimum wages are even closer to zero. That is, a monopsony economy matching the same aggregates provides an even weaker case for minimum wages.

Internally calibrated parameters are in Table 1B. 'Shifters', $\widetilde{Z}$ and $\overline{\varphi}$, are pinned down exactly by average firm size and the fraction of workers that earn below \$15 per hour. The average size of a firm at the commuting zone level is 22.83 (LBD), and 30 percent of workers earn below \$15 per hour (CPS). We assume productivity is log normally distributed. The standard deviation σ_z and decreasing returns α are identified by the average level of concentration in labor markets, and the labor share.[12] Our inferred level of productivity dispersion ($\sigma_{\log z} = 0.31$) is slightly less than direct empirical estimates.[13] We infer moderate decreasing returns ($\alpha = 0.94$), which implies a relatively elastic marginal revenue product of labor, hence firms shrink quickly in Region III. The capital share, which we set to 0.18 (Barkai, 2020), determines γ.

3.2 Optimal $\underline{w}$ with homogeneous workers

To compute optimal policy, we rely on the *consumption equivalent welfare gain relative to a no minimum wage economy* (henceforth, *welfare gains*). This is the proportional increase in consumption $\Lambda(\underline{w})$ that delivers the same utility as the minimum wage economy.

$$\text{Definition of } \Lambda(\underline{w}): \qquad U\Big(\Big(1+\Lambda(\underline{w})\Big)C(0), N(0)\Big) = U\Big(C(\underline{w}), N(\underline{w})\Big).$$

We find that the possible welfare gains are small. Figure 4A shows that $\Lambda(\underline{w})$ attains a maximum of 0.22% at \$7.65. A counterfactual economy in which we keep $\underline{w}$ at zero and increase TFP $\widetilde{Z}$ by 0.22% attains the same welfare gain. That these coincide provides a strong justification of our welfare metric.

That welfare gains are small is not because there are none to be had. A counterfactual that sets $\mu_{ij} = 1$ delivers the efficient allocation and yields a welfare gain of 6.3%. Welfare gains are only 3% of those attainable from removing labor market power, which has been a stated aim of minimum wage policy.

Figure 4B decomposes welfare into the component associated with misallocation $\omega(\underline{w})$, and shadow markdowns $\widetilde{\mu}(\underline{w})$, by feeding each into the economy separately. At the optimal minimum wage, the gain is evenly split. With an employment weighted average markdown of 0.72, markdowns have room to improve and are still improving at \$7.65. However, at higher minimum wages, the negative

[12] More productivity dispersion increases the market power of the most productive firms. This *increases concentration and decreases the labor share*. More linear technology also makes the most productive firms larger, but reduces profits. This *increases concentration and increases the labor share*.

[13] Decker, Haltiwanger, Jarmin, and Miranda (2020) derive establishment-level TFP following production function estimation at the 3-digit NAICS level for 2000 to 2013. They then compute the average of *within-6-digit-industry standard deviation of log TFP* and obtain 0.38 (their Figure 3A) and in a narrower industry classification than our baseline. BLS *Dispersion Statistics on Productivity* computes average within-4-digit-industry log interquartile range (i.e. $IQR = \log\left(z(p75)/z(p25)\right)$) of TFP over 2012-2017 between 0.45 (Chart 4) and 0.55 (Chart 3), depending on weighting. In our model, this statistic is 0.42 at the 3-digit level, where one would expect greater dispersion.

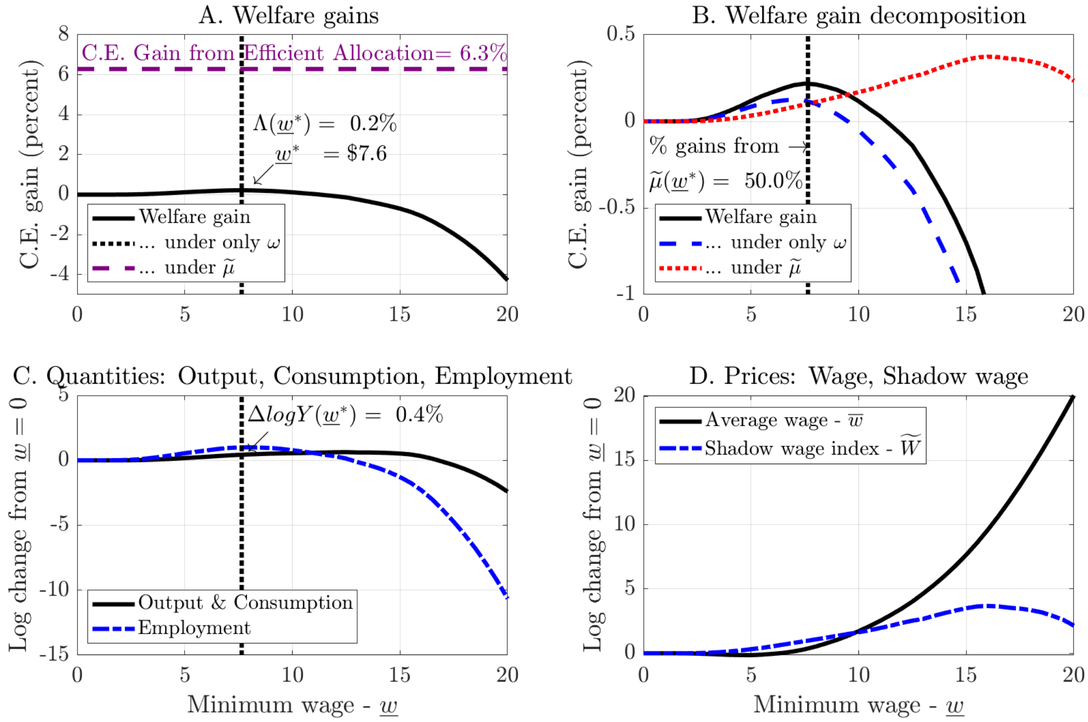

Figure 4: Minimum wages and welfare

Notes: In all cases we plot objects from the equilibrium under various values of the minimum wage $\underline{w}$, on the horizontal axis. In all cases the vertical axis plots differences from a zero minimum wage economy. **Panel A.** Plots the consumption equivalent welfare gains: $\Lambda(\underline{w})$. The long-dash purple line illustrates the welfare gain from the competitive allocation. The solid black line illustrates the welfare gain from the minimum wage in the monopsony economy, $\Lambda(w)$ defined in the text. **Panel B.** Plots the consumption equivalent welfare gains due to markdowns and misallocation. The long-dash blue line illustrates the welfare gain $\Lambda(w)$ resulting from changes in allocative efficiency ω_k only. The dotted red line illustrates the welfare gain $\Lambda(w)$ resulting from changes in markdowns $\widetilde{\mu}_k$ only. **Panel C.** Plots the percent change in output (which equals the percent change in consumption; solid) and employment (bodies; dashed). Note that employment is measured in total units of labor $\int \sum_j n_{ij} dj$, rather than the disutility term. **Panel D.** Plots the shadow wages index (dashed) and average wage (solid).

forces discussed in Figure 3B dominate. Misallocation worsens as employment is diverted from the most productive firms, sharply deteriorating welfare.

Output, consumption and employment. Aggregate employment, output and consumption have small gains that also deteriorate quickly at higher minimum wages (Figure 4C). The small output gains track the small efficiency gains. At the optimal minimum wage of $7.65, output gains reach a mere 0.40%. The profile of these aggregates will be similar when we include household heterogeneity in Section 4. Since these aggregates track the value-added in production, rather than the distribution of resources, the efficiency implications of the minimum wage will also be similar.

Wages. The average wage increases monotonically with the minimum wage, however the path of aggregate employment is hump-shaped (Figure 4D). Aggregate employment does not follow the average wage, since the average wage no longer captures market forces of supply and demand. The aggregate shadow wage $\widetilde{W}$, however, does represent the market clearing price for labor. It increases as markdowns narrow, and then falls as shadow markdowns widen, encoding binding rationing constraints at Region III firms. In the aggregate, employment follows the shadow wage index. A direct implication for empirical research is to reduce emphasis on the response of wages to minimum wage laws, since wages themselves are not welfare relevant.

3.3 Mechanisms

Two questions arise: (i) why are efficiency gains small?, (ii) what economic forces lead the gains to be positive? We shed light on both questions below.

3.3.1 Why are efficiency gains small? - Firm heterogeneity mutes *Direct effects*

Its well-known since Robinson (1933) that a minimum wage can completely offset the efficiency losses due to the market power of a solo monopsonist by setting the minimum wage equal to the perfectly competitive wage. Is a national or market minimum wage in the presence of realistic firm heterogeneity just as effective? No.

Efficiency gains are small for five main reasons. First, the minimum wage binds first at low productivity firms. Second, low productivity firms have a small share of employment and narrow markdowns. Third, the direct monopsony channel operates in a narrow window due to narrow markdowns and elastic labor supply. Fourth, because firm labor demand is elastic, gains quickly become losses as firms shrink beyond competitive levels of employment. Finally, the spillover channel is quantitatively limited: increases in the minimum wage do not notably affect the employment choices of the largest firms.

To gain intuition, Figure 5 provides an illustrative example of a market with two firms: a less productive Corner store and a more productive Supermarket. Both have monopsony power. The faded lines in Figure 5A correspond to equilibrium employment, wages and markdowns for each firm under $\underline{w} = 0$. We point out how features of the data would inform a comparison of two such firms. First, our calibration implies the variation across the firms in size is substantial. There are on average 113 firms in each market. But the average HHI is 0.11. This is what one would observe from a market with around 10 equally sized firms. To match this our calibration requires dispersion in productivity ($\sigma_z = 0.31$), which is in line with empirical estimates.[14,15] Second, these differences imply substantially wider markdowns at the Supermarket. Being much smaller, the Corner store faces more elastic labor and has a narrow markdown near 1. The Supermarket has a wide markdown. Third, high concentration implies the Supermarket has a large share of employment, and hence *overall efficiency losses* are driven by its markdown.

With these features in mind, suppose the government follows Robinson (1933) and sets a minimum wage equal to the competitive wage of the Corner store (solid lines of Panel A). In partial equilibrium, this doesn't effect the Supermarket and removes the efficiency loss induced at the Corner store. Market employment increases. But, because the Corner store's markdown is small and the Supermarket is unaffected, this *Direct* effect is small.

This intuition extends to markets with many more firms. The window of productivity for which firms like the Corner Store are in Region II is narrow: low productivity firms with small market shares face elastic labor supply curves ($\eta = 10$). Small increases in $\underline{w}$ quickly increase their employment to the competitive level, beyond which they ration workers. The numerical example in Figure 6 demonstrates this point in a market that we randomly draw from the set of markets with 200 firms, imposing $\underline{w}$ of \$15. Only a small set of firms are in Region II (each represented by a diamond). The line in Panel B shows the efficient level of employment for each firm when markdowns are all equal to 1. Note that even medium productivity Region I firms have employment close to the competitive level. The efficiency losses only emerge at very large firms in Region I.

[14] Alternatively, productivity differences could be smaller, but α or η could be higher. We already have α close to constant returns. We already have η equal to 10.85.

[15] We also match the empirical *size-wage-elasticity*. Pooling data from all markets, and regressing log average wage on log employment we obtain a coefficient of 0.05 which lies between the estimates in Bloom, Guvenen, Smith, Song, and von Wachter (2018) (see their Figure 1 which reports size-wage elasticities between 0.04 and 0.06).

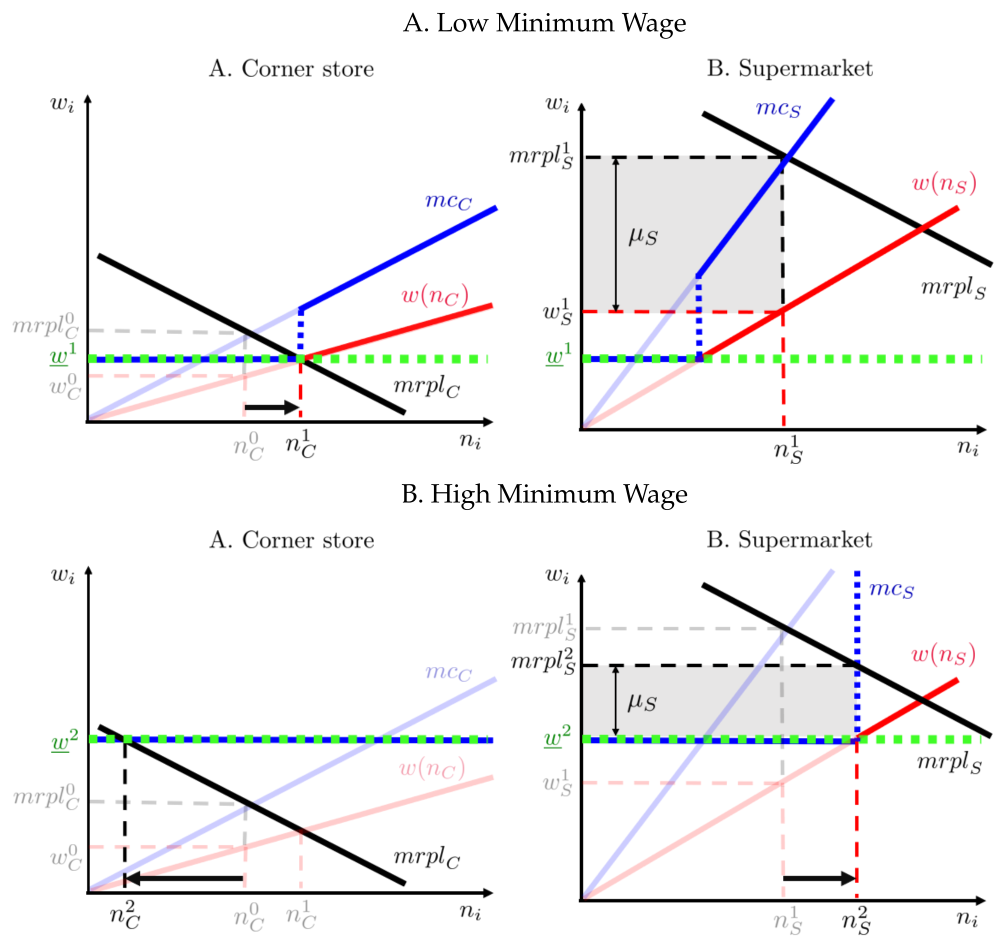

Figure 5: Productivity heterogeneity limits efficiency gain from minimum wage

What if the government raises $\underline{w}$ to target the efficiency losses at these larger firms, like the Supermarket? Figure 5B shows that eating into the Supermarket's efficiency losses comes at the cost of rationing the employment at the Corner store. We estimate a relatively elastic marginal revenue product of labor ($\alpha = 0.94$). Employment is therefore rationed quickly at the Corner store as soon as the minimum wage is set too high. Crosses (red) in Figure 6 extend this logic to our multi-firm numerical example. The widening gap between each firm and competitive employment shows severe rationing of employment at low productivity firms.

The above arguments are driven by the significant amount of firm heterogeneity in the data. Interestingly, we find that overall efficiency gains are still small (though the efficiency maximizing minimum wage is significantly higher), when there is much less firm heterogeneity. We show this by simulating a model econ-

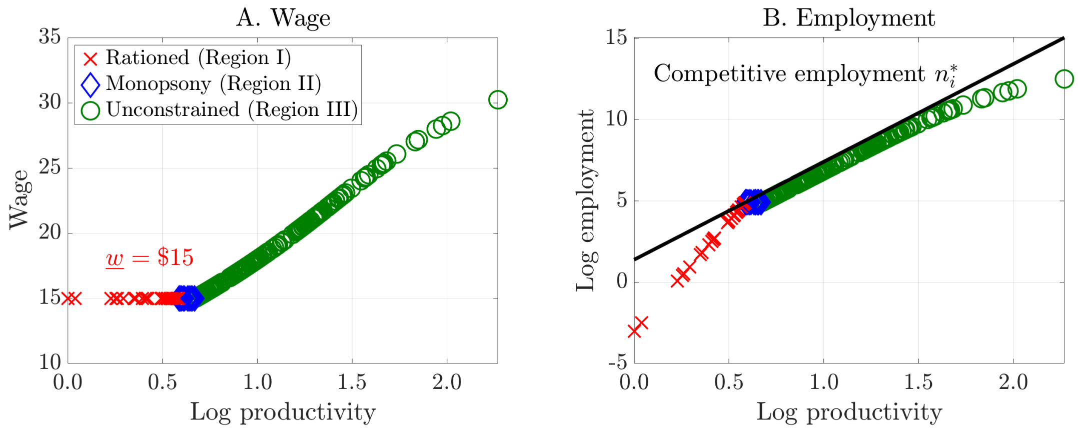

Figure 6: Small efficiency gains from minimum wage in a 200 firm market

Notes: This figure is produced using parameters from Table 1. We impose $\underline{w}$ = \$15 and solve for the new general equilibrium allocation. We then isolate one single market with $M_j = 200$ and plot the corresponding allocations. Crosses $\times$'s (red) are Region I firms, diamonds (blue) are Region II firms, and circles (green) are Region III firms. The solid line (black) represents competitive employment, where we fix market $(n_j, \widetilde{w}_j)$ and solve out firm labor supply $n_{ij} = (w_{ij}/w_j)^{\eta} n_j$ and demand under $\widetilde{\mu}_{ij} = 1$: $w_{ij} = \widetilde{\alpha}\widetilde{z}_{ij} n_{ij}^{\widetilde{\alpha}-1}$.

omy with half the productivity dispersion of our baseline calibration. With less productivity dispersion, markets are counterfactually less concentrated: the average HHI is 0.06 versus 0.11 in the data. With less productivity dispersion, the optimal minimum wage is \$10.60, approximately \$3 dollars higher than the baseline (Figure 7A). However, welfare and output gains double but remain quantitatively small: welfare increases by 0.5% (baseline: 0.2%) and output increases by 1.1% (baseline: 0.4%). Minimum wages yield small efficiency gains even with counterfactually low productivity heterogeneity. [16]

The final reason the efficiency gains are small is even though our model matches empirical evidence on spillovers across workers (Section 7), an increase in the minimum wage has quantitatively negligible spillovers on the markdowns of high productivity, unconstrained firms. These firms are shown by the circles in Figure 6, and are responsible for the majority of the departure from competitive employment. They respond little to the increase in wages of their low wage competitors, as their low wage competitors have small market shares.

[16] In further sensitivity analysis in Appendix A.J, we find that decreasing (*increasing*) elasticities (θ, η) by 30% increases (*decreases*) the optimal minimum wage by 70c (*30c*), and leaves welfare gains almost unchanged.

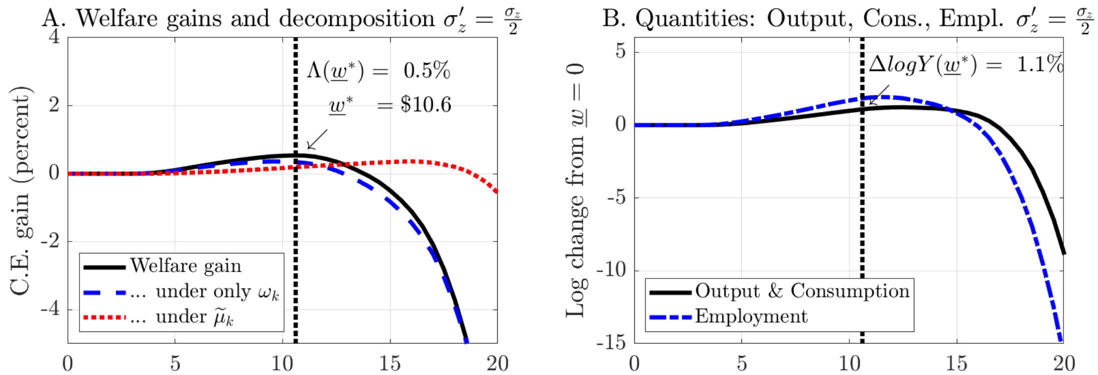

Figure 7: Minimum wages and welfare - Half dispersion in productivity

Notes: This figure computes the optimal minimum wage when productivity dispersion is halved $\sigma_z' = \frac{\sigma_z}{2}$. See notes to Figure 4. **Panel A.** Plots the consumption equivalent welfare gains of each household: $\Lambda(\underline{w})$. The solid black line illustrates the welfare gain from the minimum wage in the monopsony economy, $\Lambda(w)$ defined in the text. The long-dash (blue) line illustrates the welfare gain $\Lambda(w)$ resulting from changes in allocative efficiency ω_k only. The dotted (red) line illustrates the welfare gain $\Lambda(w)$ resulting from changes in markdowns $\widetilde{\mu}_k$ only. **Panel B.** Plots the percent change in output (which equals the percent change in consumption) and employment (bodies).

3.3.2 What does account for the positive gains? - *Reallocation* and *Spillovers*

Figure 4B demonstrated that the small positive efficiency gains from the minimum wages are equally attributable to positive reallocation (ω) and narrower markdowns ($\widetilde{\mu}$). We argue that the within-market *Reallocation* and *Spillover* channels, which are present in markets with a finite number of oligopsonists under $\eta > \theta$, are crucial for capturing the (small) efficiency benefits, not *Direct* effects.

We separate the importance of *Direct* effects versus *Spillovers* and *Reallocation* by comparing our baseline economy to an economy in which $\eta = \theta$. This is the frequently used *monopsonistically competitive model* with firm heterogeneity. *Direct* effects are present but *Spillovers* and *Reallocation* are not.[17] To compare models, we set $\eta = \theta = 3.02$. This gives the same aggregate labor share as the baseline economy, and hence the same scope for *Direct* effects.[18] In fact, since markdowns are now wider at small firms, this gives *Direct* effects an even better shot.

RE

Table 2 shows that in a monopsonistically competitive economy, the efficiency maximizing minimum wage is only \$0.70. Absent positive effects of *Spillovers* and

[17]When $\theta = \eta$ all firms in all markets are effectively infinitesimal in a national labor market, with no distinction between local labor markets. When a small firms enters Region III, employment is reallocated into the aggregate pool of labor, rather than up the ladder within the market. Hence we refer to this as no *Reallocation* effects in the way we discussed previously.

[18]We recalibrate 'shifters', $\{\overline{\varphi}_h, \xi_h, \kappa_h\}_{h=1}^{H}$, $\widetilde{Z}$, $\overline{\varphi}$, to match the same moments in Appendix O, Table A1.

	Optimal $\underline{w}$
Baseline - Granular firms in local markets - *Oligopsony* - $\eta > \theta$	\$7.65
Alternative - Infinitesimal firms in national market - *Monopsony* - $\eta = \theta$	\$0.70

Table 2: Minimum wages and welfare - Role of granular markets

Reallocation welfare gains are almost completely shut down. Firm heterogeneity severely limits *Direct* effects to the point where they are unable to improve welfare.

The positive reallocation improvements resulting from minimum wage hikes in Figure 4B come from workers moving up the local job ladder. In our baseline, jobs lost at Region III firms are mostly reallocated *within-market* to local firms with higher productivity (recall Figure 2B). As local labor markets are granular, these firms have discretely higher productivity. In a monopsonistically competitive economy, when a small firm shrinks in Region III, their employment is reallocated into the aggregate pool of labor N, rather than up the ladder within the market into n_j.

Likewise, the positive markdown improvements resulting from minimum wage hikes in Figure 4B are due to *Spillovers*, not *Direct effects*. As smaller, less productive firms raise their wages, larger, more productive firms also increase wages due to strategic complementarities, i.e. spillovers. While these effects are small, they yield a motive for positive minimum wages which is absent from the non-strategic model.

4 Heterogeneous workers

We generalize our economy to include H heterogeneous households indexed by $h \in \{1, \ldots, H\}$. Our main result is the following: once we adjust for the redistributive effects of minimum wages, efficiency gains are as small as in the homogeneous household case, and the optimal minimum wage is effectively unchanged. This section is intentionally terse, since most details follow from the prior section (the prior model is nested).[19]

Agents. Households differ in their measure π_h, disutility of labor $\overline{\varphi}_h$, labor productivity ξ_h and share of aggregate non-labor income κ_h.

[19] All derivations and definition of equilibrium can be found in Appendix A.G.

Goods and technology. Firms use capital and labor of each type n_{ijh}. Firm-ij produces y_{ij} units of net-output according to

$$y_{ij} = \overline{Z} z_{ij} \sum_{h=1}^{H} \left(\left[\xi_h n_{ijh} \right]^{\gamma} k_{ijh}^{1-\gamma} \right)^{\alpha} \quad , \quad \gamma \in (0,1] \quad , \quad \alpha > 0$$

where k_{ijh} is capital allocated to worker type h. Production has a unit elasticity of substitution between capital and labor of each type. While a range of estimates of the elasticity of substitution between capital and labor are reported in empirical papers, many find elasticities in the range of 0.7 to 1.2 (see Section 7 for discussion). The labor-labor elasticity of substitution between types h and h' is

$$\rho\left(h, h'\right) := -\frac{d \log \left(n_{ijh'} / n_{ijh}\right)}{d \log MRTS(h, h')} = \frac{1-(1-\gamma)\,\alpha}{1-\alpha} \quad , \quad MRTS(h, h') = \frac{dy_{ij}/dn_{ijh}}{dy_{ij}/dn_{ijh'}}. \tag{9}$$

In Appendix O (Section C.5), we vary α to provide robustness of our main results with respect to the degree of substitutability across labor types.[20]

Household problem. Each household has concave preferences over per-capita consumption and disutility from supplying labor:

$$\mathcal{U}_h = \sum_{t=0}^{\infty} \beta^t u^h \left(\frac{c_{ht}}{\pi_h}, n_{ht} \right) = \sum_{t=0}^{\infty} \beta^t \left[\frac{\left(c_{ht}/\pi_h\right)^{1-\sigma}}{1-\sigma} - \frac{1}{\widetilde{\varphi}_h^{1/\varphi}} \frac{n_{ht}^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} \right]. \tag{10}$$

The type-specific labor supply index n_{ht} is a nested-CES over markets and firms:[21]

$$n_{ht} := \left[\int_0^1 n_{jht}^{\frac{\theta+1}{\theta}} \, dj \right]^{\frac{\theta}{\theta+1}} \quad , \quad n_{jht} := \left[\sum_{i=1}^{M_j} n_{ijht}^{\frac{\eta+1}{\eta}} \right]^{\frac{\eta}{\eta+1}},$$

Household h has its own budget constraint. This means that within-household risk associated with labor being rationed due to the minimum wage is insured, but across-household risk is not. We discuss this further in Section 7. Endowments of initial capital $\{k_{k0}\}$ are a free-parameter of the competitive equilibrium. We assume each household's share of initial K_0 is equal to its share of profits:

$$c_{ht} + k_{ht+1} = \int \sum_{i=1}^{M_j} w_{ijht} n_{ijht} \, dj + R_t k_{ht} + (1-\delta) k_{ht} + \kappa_h \Pi_t \quad , \quad k_{h0} = \kappa_h K_0. \tag{11}$$

Given all prices, household h chooses n_{ijht} and k_{ht+1} to maximize utility (10) subject to (11) and labor rationing constraints, $n_{ijht} \leq \overline{n}_{ijht}$.

[20] What do we miss by not having a CES formulation? Simply that our production function is homogeneous of degree $\gamma\alpha$ in the vector $\mathbf{n}_{ij}$, not one. But this is without loss given we want to keep decreasing returns as per our theoretical exercises in Section 1.

[21] The parameter $\widetilde{\varphi}_h$ expresses the disutility of labor supply on a per capita basis which we normalize by an aggregate measure $\overline{\varphi}$: $\widetilde{\varphi}_h = \left(\overline{\varphi}_h / \overline{\varphi}\right) \pi_h^{1+\varphi}$

First order conditions can again be rewritten in terms of shadow wages, with indices defined by household type h. As before, the household shadow wage index $\widetilde{w}_{ht}$ determines the allocation of labor n_{ht}:

$$n_{ijht} = \left(\frac{\widetilde{w}_{ijht}}{\widetilde{w}_{jht}}\right)^{\eta}\left(\frac{\widetilde{w}_{jht}}{\widetilde{w}_{ht}}\right)^{\theta} n_{ht} \qquad , \qquad n_{ht} = \pi_h \widetilde{\varphi}_h \widetilde{w}_{ht}^{\varphi}\left(\frac{c_{ht}}{\pi_h}\right)^{-\sigma\varphi} \tag{12}$$

Unlike the homogeneous worker economy, aggregate capital income and profits link households through wealth effects on labor supply (via c_h).[22]

Firm problem. At a particular allocation and prices, a firm's profits are:

$$\pi_{ijt} = \overline{Z} z_{ijt} \sum_{h=1}^{H}\left(\left[\xi_h n_{ijht}\right]^{\gamma} k_{ijht}^{1-\gamma}\right)^{\alpha} - R_t \sum_{h=1}^{H} k_{ijht} - \sum_{h=1}^{H} w_{ijht} n_{ijht} \tag{13}$$

The firm's problem is to choose $(n_{ijht}, \overline{n}_{ijht}, w_{ijht}, k_{ijht})$ for each h in order to maximize profits (13), subject to each household's labor supply schedule (12), the rationing constraint $n_{ijht} \leq \overline{n}_{ijht}$, and the minimum wage $w_{ijht} \geq \underline{w}$.

Since profits are additively separable across household types $h = 1, \ldots, H$, the firm solves each problem separately, choosing $(n_{ijht}, \overline{n}_{ijht}, w_{ijht}, k_{ijht})$ as per the firm in the homogeneous worker economy. The firm's optimal rationing constraint is still determined by the level of labor at which the firm's marginal revenue product of labor is equal to the minimum wage, e.g. $mrpl(\overline{n}_{ijht}) = \underline{w}$. Optimizing out the choice of type-h capital from the above, the firm's profits for type-h labor are

$$\pi_{ijht} = \widetilde{Z}\widetilde{z}_{ijt}\widetilde{\xi}_h n_{ijht}^{\widetilde{\alpha}} - w_{ijht} n_{ijht} \ , \ \widetilde{Z} := \overline{Z}^{\frac{1}{1-(1-\gamma)\alpha}} \ , \ \widetilde{\xi}_h := \xi_h^{\widetilde{\alpha}} \ , \ \widetilde{\alpha} := \frac{\gamma\alpha}{1-(1-\gamma)\alpha}.$$

Hence the weakly optimal rationing constraint $\overline{n}_{ijht}$ satisfies

$$\underline{w} = \widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ijt}\widetilde{\xi}_h \overline{n}_{ijht}^{\widetilde{\alpha}-1} \ , \ \overline{n}_{ijht} = \left(\frac{\widetilde{\alpha}\widetilde{Z}\widetilde{\xi}_h\widetilde{z}_{ijt}}{\underline{w}}\right)^{\frac{1}{1-\widetilde{\alpha}}} \ , \ \widetilde{z}_{ijt} := \left[1-(1-\gamma)\alpha\right]\left(\frac{(1-\gamma)\alpha}{R_t}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}} z_{ijt}^{\frac{1}{1-(1-\gamma)\alpha}}.$$

The definition of equilibrium and firms' optimal employment, wage, and rationing constraints follow directly from Section 1. Likewise, firms can be split into Regions I, II, and III using identical definitions as Section 1.

Aggregation. As before, the economy can be aggregated at the household level exploiting a household level shadow markdown $\widetilde{\mu}_h$ and misallocation ω_h. Labor supply, labor demand and output are then pinned down by these endogenous

[22] For type-h, steady-state capital income is $\kappa_h((R-\delta)K + \Pi)$. Aggregate capital demand is $K = \alpha(1-\gamma)Y/R$, which clears at the initial capital stock under $1 = \beta(R + (1-\delta))$. Aggregate profits are $\Pi = Y - \sum_h \left[\int \sum_i w_{ijh} n_{ijh} dj\right] - RK$. Thus, aggregate capital income and profits link households via wealth effects on labor supply.

wedges, omitting time subscripts for ease of exposition:

$$n_h = \pi_h \widetilde{\varphi}_h \widetilde{w}_h^{\varphi} c_h^{-\sigma\varphi} \quad , \quad \widetilde{w}_h = \widetilde{\mu}_h \widetilde{\alpha} \widetilde{Z} \widetilde{\xi}_h \widetilde{z}_h n_h^{\widetilde{\alpha}-1} \quad , \quad y_h = \frac{1}{1-(1-\gamma)\alpha} \omega_h \widetilde{Z} \widetilde{\xi}_h \widetilde{z}_h n_h^{\widetilde{\alpha}}.$$

As before, the set of wedges $\{\widetilde{\mu}_h, \omega_h\}_{h=1}^{H}$ summarize deviations from efficiency due to labor market power *and* the minimum wage. For each household type h, shadow markdowns are captured by $\widetilde{\mu}_h$ and misallocation is captured by ω_h. This allows us to separate the efficiency effects of minimum wages into shadow markdowns and misallocation for each household type h. In the efficient allocation $\widetilde{\mu}_h = 1$ and $\omega_h = 1$ for all h.

4.1 Calibration

Data sources used to calibrate the heterogeneous worker economy are identical to the homogeneous worker economy in Section 3, with the addition of the 2016 and 2019 Survey of Consumer Finances (SCF) to discipline capital ownership.

Households. We construct twelve household types: $H = 12$. First, we split households into three education groups: those with less than a high-school diploma (NHS), those with only a high school diploma (HS), and those who have completed college. Second, we partition NHS and HS groups into five wage quintiles each.[23] Third, we split college households: those for which capital income is more than half of their wage income, whom we call owners (O), and the remainder whose primary earnings source is labor income, whom we call college workers (C).

We use the SCF to identify business owners. We measure capital income as interest and dividend income, business and farm income, and realized capital gains.[24] For 7% of the SCF, capital income is more than half of labor income. We treat all such individuals as college educated business owners (O).[25]

Model inversion. We first take population shares π_h from the CPS. Parameters that are heterogeneous across households are relative shifters in productivity and

[23] Hurst, Kehoe, Pastorino, and Winberry (2022) use a similar procedure.

[24] We also consider an alternative approach, where we determine capital income as a residual in the household budget constraint. By this approach capital income is defined as total income minus labor income and transfers. This yields a very similar split of households.

[25] When aggregated, non-college workers' capital income is not zero, but it is small, and hence our assumption that only college households are owners is reasonable. Of the households that earn more than half of their income from capital income, 80% of capital income accrues to college educated workers.

Parameters		**NHS**	**HS**	**C**	**O**
Relative population (%)	$\pi_h / \sum \pi_h$	13.2	53.7	26.1	7.0
Relative disutility labor supply	$\overline{\varphi}_h^{-\varphi} / \overline{\varphi}_C^{-\varphi}$	10.75	2.21	1.00	0.53
Relative productivity	ξ_h	0.25	0.49	1.00	0.89
Capital income share (%)	κ_h	0.10	1.64	4.30	93.96
Labor disutility shifter	$\overline{\varphi}$	— 5.05×10^6 —			
Productivity shifter	$\widetilde{Z}$	—— 16.84 ——			

Table 3: Parameters

	Model				Data			
	Non-HS	**HS**	**Coll**	**Own**	**Non-HS**	**HS**	**Coll**	**Own**
Population shares* (CPS, %)	13.2	53.7	26.1	7.0	13.2	53.7	26.1	7.0
Share of agg. labor income* (CPS and SCF, %)	3.0	38.5	46.2	12.4	3.0	38.5	46.2	12.4
Ave. earnings per hour*, (CPS, C=1)	0.40	0.59	1.00		0.40	0.59	1.00	
Capital income to labor income* (SCF)	0.01	0.02	0.05	4.00	0.01	0.02	0.05	2.62
Binding at \$15, by type (CPS, %)	76.8	45.5	10.8		68.7	38.1	11.1	
Binding at \$15, all* (CPS, %)	—— 30.6 ——				—— 30.6 ——			
Average firm size* (LBD)	—— 22.8 ——				—— 22.8 ——			

Table 4: Model versus data moments (* denotes moments that are targeted)

Notes: For Non-HS and HS household types, this table gives moments computed when aggregating across all five of the associated types of each household. This is only for presentation purposes.

labor supply disutility $\{\xi_h, \widetilde{\varphi}_h\}_{h=1}^{H}$, where $\widetilde{\varphi}_h = (\overline{\varphi}_h / \overline{\varphi})\, \pi_h^{1+\varphi}$, and shares of aggregate profits and capital income $\{\kappa_h\}_{h=1}^{H}$.

We normalize $\xi_h = \widetilde{\varphi}_h = 1$ for college worker households. For any $\{\kappa_h\}_{h=1}^{H}$, the remaining productivity and labor disutility parameters can be inverted from data on relative average labor earnings per hour and households' shares of aggregate labor income, which we compute in the CPS. For example, relative productivities $\{\xi_h\}_{h=1}^{H}$ are inverted so that the average wage of non-high-school (high-school) workers is 40 percent (59 percent) of the average college wage.[26] Relative disutilities of labor supply $\{\widetilde{\varphi}_h\}_{h=1}^{H}$ are pinned down by shares of total labor income.

We choose $\{\kappa_h\}_{h=1}^{H}$ for each of the eleven non-owner households to exactly match their empirical ratio of total capital income to total labor income, measured in the SCF. This is less than 0.05% for all non-owner households, providing further support for our approach of including owners as a separate group. Owners' share of capital income is a residual.

As in the homogeneous worker economy, common parameters $\widetilde{Z}$ and $\overline{\varphi}$ are

[26] We assign college worker and owner households the same wage. This allows us to combine SCF and CPS data since we do not observe assets in the CPS. In the SCF, labor earnings are similar across the two college household types.

inverted to exactly match average firm size (22.8) and fraction of workers that earn below \$15 per hour (30 percent). It does well on the non-targeted fraction of college workers below \$15 (10.8% vs. 11.1% in data), high school workers (45.5% vs. 38.1% in data), and non-high school workers (76.8% vs. 68.7% in data).

Tables 3 and 4 report averages of parameters and aggregated moments for the four broad household groups. Parameters and moments for all 12 households are reported in Appendix O.A.

5 Optimal $\underline{w}$ with heterogeneous workers

We first compute the welfare maximizing minimum wage under Utilitarian welfare weights. We then separate welfare gains into an efficiency component and a welfare-weight-dependent redistribution component using elements of Floden (2001) and Dávila and Schaab (2022).

Measuring Welfare. Under a minimum wage $\underline{w}$, we compute each household's *consumption equivalent welfare gain relative to a no minimum wage economy* (henceforth, *welfare gains*) as the proportional increase in consumption $\lambda_h(\underline{w})$ that delivers the same utility as the minimum wage economy.[27] We define the Utilitarian welfare gain, $\Lambda_\pi(\underline{w})$, which values households in accordance to their population share π_h.

$$\text{Definition of } \lambda_h(\underline{w}): \quad u^h\left(\left(1+\lambda_h(\underline{w})\right)\frac{c_h(0)}{\pi_h}, n_h(0)\right) = u^h\left(\frac{c_h(\underline{w})}{\pi_h}, n_h(\underline{w})\right)$$

$$\text{Definition of } \Lambda_\pi(\underline{w}): \quad \sum_h \pi_h u^h\left(\left(1+\Lambda_\pi(\underline{w})\right)\frac{c_h(0)}{\pi_h}, n_h(0)\right) = \sum_h \pi_h u^h\left(\frac{c_h(\underline{w})}{\pi_h}, n_h(\underline{w})\right).$$

With power utility, $\Lambda_\pi(\underline{w})$ is a harmonic mean of the $\lambda_h(\underline{w})$'s, with weights given by a transformation of π_h's.

5.1 Results

Figure 8 depicts the optimal minimum wage under a Utilitarian welfare criteria. Panel A shows that the Utilitarian welfare maximizing minimum wage is \$11.00. At a minimum wage of \$11.00, the Utilitarian welfare gain is of 2.8% of consumption. Panel A also plots welfare gains from the efficient allocation ($\mu_{ijh} = 1, \ \forall ijh$).

[27] The choice to benchmark our welfare gains relative to an economy with a zero minimum wage is easy to amend and has little bearing on our results.

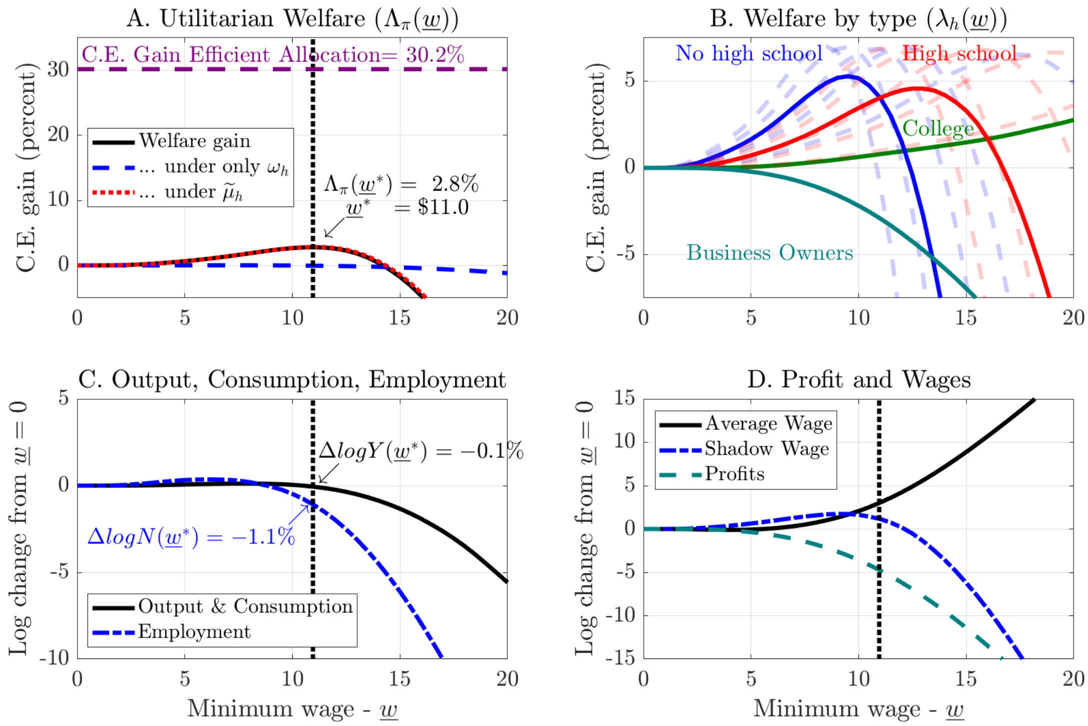

Figure 8: Minimum wages and welfare

Notes: In all cases we plot objects from the equilibrium under various values of the minimum wage $\underline{w}$, on the horizontal axis. In all cases the vertical axis plots differences from a zero minimum wage economy. **Panel A.** Plots the aggregate consumption equivalent welfare gains $\Lambda_\pi(\underline{w})$ (solid line, black) attributable to markdowns (dotted line, red line) and misallocation (dashed line, blue). The efficient allocation welfare gains are denoted by the upper horizontal dasehd line (purple) and is obtained by setting $\mu_{ijh} = 1 \forall \{ijh\}$. The optimal minimum wage is the black dashed vertical line. **Panel B.** Plots the consumption equivalent welfare gains $\lambda_h(\underline{w})$ for non-highschool workers, high school workers, college workers, and business owners. **Panel C.** Plots the log change in output and consumption (which are equivalent) and the change in employment (measured in bodies). The optimal minimum wage is the dashed vertical line. **Panel D.** Plots average wages (solid line, black), the average wage index across worker types (dash-dot line, blue), and business profits (dashed line, teal).

The consumption equivalent gain to a Utilitarian planner from the efficient allocation is 30.2%. Thus even when redistributive gains are included, the optimal minimum wage captures less than one-tenth of the potential gains from the efficient allocation. In contrast to the homogeneous worker case, Utilitarian gains are primarily driven by narrower markdowns. Narrowing markdowns directly raise wages of households a Utilitarian planner cares about. Resolving misallocation is of little value to a planner who cares about redistribution.

Welfare is hump-shaped for each worker type but with different welfare maximizing minimum wages (Panel B). For minimum wages up to $10 dollars, all worker types are better off, except for business owners who are hurt by lower profits. What drives the worker welfare gains? The next panels establish that the gains are driven almost entirely by a redistribution of business profits to households.

Output and employment are effectively non-increasing in the minimum wage (Panel C). With no additional final goods being produced, welfare gains must stem from redistribution. Eventually, for high enough minimum wages, there are severe output and employment losses. In fact, at the Utilitarian optimum, production is 0.1% lower and employment is 1.1% lower.

Despite this, average wages monotonically increase and profits monotonically decline (Panel D). Shadow wages also initially increase. However, similar to the homogeneous worker case, shadow wages sharply fall beyond a minimum wage of \$12 as employment rationing becomes severe. These wage gains ultimately drive the worker welfare improvements observed in Panel B and since production does not increase, these wage gains are a pure transfer from business owners to households. Not shown here, the labor share monotonically increases.

5.2 Efficiency maximizing minimum wage

A serious drawback of the above results is that they ultimately depend on the particular choice of social welfare weights that the household gains $\lambda_h(\underline{w})$ are integrated over. To deal with this issue we parse welfare gains into an efficiency component, which reflects gains from greater aggregate consumption and employment, and a redistribution component, which reflects welfare-weight-dependent gains from reallocating resources.

We first define social welfare $\mathcal{W}$ and normalized social welfare $\mathcal{W}_\Gamma$ as follows:

$$\mathcal{W} := \sum_h \pi_h u^h \left(\frac{c_h}{\pi_h}, n_h\right), \ \ \mathcal{W}_\Gamma := \frac{\mathcal{W}}{\Gamma}, \ \ \Gamma := \sum_h \pi_h u_c^h \left(\frac{c_h}{\pi_h}, n_h\right) \frac{c_h}{\pi_h} = \sum_h \pi_h \left(\frac{c_h}{\pi_h}\right)^{1-\sigma}.$$

Here, Γ converts utils into consumption equivalent terms (Dávila and Schaab, 2022). To a first order, dividing by marginal utility converts welfare into consumption units; further dividing by consumption converts it into percentage deviations. Unlike Dávila and Schaab (2022) we do not take a first-order approximation of the welfare function.

We then apply the same logic as Floden (2001) to isolate the aggregate efficiency component of welfare. Define aggregate consumption and employment ($C = \sum_h c_h$, $N = \sum_h n_h$, where N and n_h are employment indices), and households' shares ($s_h^C = c_h/C$, $s_h^N = n_h/N$). Take any counterfactual allocation denoted with primes (e.g. c_h'). Normalized welfare gains are the sum of *aggregate efficiency* (AE) and *redistribution* (RE) gains:

$$\underbrace{\mathcal{W}_\Gamma' - \mathcal{W}_\Gamma}_{\text{Total Welfare (TOT)}} = \underbrace{\sum_h \frac{\pi_h}{\Gamma}\left[u^h\left(\frac{s_h^C C'}{\pi_h}, s_h^N N'\right) - u^h\left(\frac{s_h^C C}{\pi_h}, s_h^N N\right)\right]}_{\text{Aggregate efficiency (AE)}} + \underbrace{\sum_h \frac{\pi_h}{\Gamma}\left[u^h\left(\frac{s_h^{C\prime} C'}{\pi_h}, s_h^{N\prime} N'\right) - u^h\left(\frac{s_h^C C'}{\pi_h}, s_h^N N'\right)\right]}_{\text{Redistribution (RE)}} \tag{14}$$

Aggregate efficiency (AE) captures the effects of C and N, holding household shares $\{s_h^C, s_h^N\}$ fixed. Gains only accrue from increasing the size of the "economic pie." Redistribution (RE) captures the effects of s_h^C and s_h^N, holding aggregates $\{C, N\}$ fixed. Gains only accrue from redistributing resources. Below we report the share of gains attributable to aggregate efficiency $\frac{AE}{TOT}$ and redistribution $\frac{RE}{TOT}$.

Figure 9 applies (14) and establishes two key findings. First, at the Utilitarian optimal minimum wage, efficiency gains are negative. Of the 2.8% welfare gains enjoyed by the Utilitarian planner, 102.5% comes from redistribution and −2.5% comes from efficiency gains. Intuitively, since less goods are being produced at the optimum (Figure 8C), the size of the "economic pie" is smaller. From the perspective of a Utilitarian planner, the minimum wage can burn resources in order to achieve some redistribution.

Second, the highest attainable efficiency gain is less than 0.10% of consumption and occurs at a minimum wage of \$7.35. These gains are less than one twentieth of the peak Utilitarian gains (2.8%). It is not a coincidence that this lies on top of our estimate for $\underline{w}^*$ in the homogeneous worker economy. The decomposition removes the redistributive motives of minimum wages which is exactly what the homogeneous worker economy accomplishes as well. Optimal minimum wages differ slightly due to the production technology difference across worker types, but the story is extremely similar: minimum wages are ineffective at reducing monopsony power.

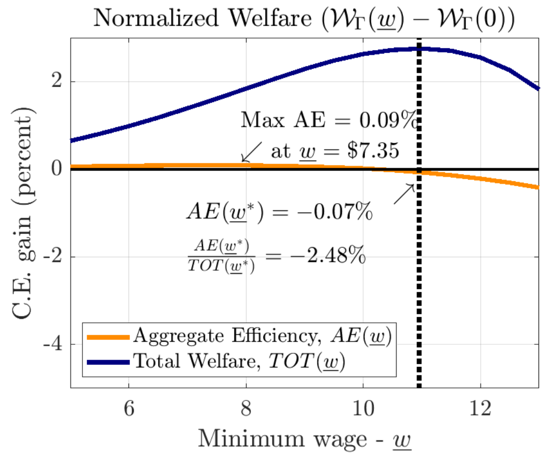

Figure 9: Minimum wages and welfare

Notes: This figure plots the normalized welfare gain $\mathcal{W}_\Gamma(\underline{w}) - \mathcal{W}_\Gamma(0)$ and the corresponding aggregate efficiency component AE of welfare. The units of both objects are consumption equivalent units. We multiply both series by 100 to express it in percent. See equation (14) and corresponding text for additional discussion.

6 Redistribution

While the efficiency gains from the minimum wage are small, the overall gains, driven by redistribution from business owners to workers, are more substantial. This section asks whether gains from redistribution under Utilitarian weights survive in a tax and transfer system that has the empirical degree of redistribution built into it. This is a pertinent question given the existence of (i) the EITC, which provides a subsidy for low income households and (ii) progressive income taxes which redistribute from high to low income individuals.

First, we find that the redistributive properties of the minimum wage are largely unaffected by existing tax and transfer policy. Second, we find that progressive taxation amplifies monopsony power, widening markdowns. Consistent with the intuition developed above, this extends Region II, providing more scope for minimum wages to increase employment. Third, we explore commonly used proxies for redistribution, including the college wage premium and wage dispersion, and discuss their suitability for guiding policy.

6.1 Taxes and transfers

We augment our model with taxes in the spirit of Benabou (2002) and Heathcote, Storesletten, and Violante (2017) (henceforth HSV). A worker of household h working at firm ij, receives after tax income $\lambda w_{ijh}^{1-\tau}$. We take $\tau = 0.181$ from HSV. The parameter λ determines the point at which subsidies becomes taxes. We choose λ to match the point at which the EITC phases out to zero.[28] Figure 10A shows that this formulation provides an excellent fit to the EITC, delivering a smooth version of the phase-in, plateau and phase-out. It then delivers progressivity over the entire tax and transfer system consistent with empirical estimates.

A novel feature of this extension is the interaction between progressive taxes and monopsony. Factorizing the rationing constraint multiplier, optimal household labor supply is:[29]

$$n_{ijh} = \left(\frac{\widetilde{w}_{ijh}}{\widetilde{w}_{jh}}\right)^{(1-\tau)\eta} \left(\frac{\widetilde{w}_{jh}}{\widetilde{W}_h}\right)^{(1-\tau)\theta} n_h. \tag{15}$$

For each increase in w_{ijh} (or $\widetilde{w}_{ijh}$), the household pays marginally higher taxes, requiring the firm to further increase wages to attract the same amount of labor. This is encoded in *lower* labor supply elasticities, scaled down by $(1-\tau)$. Internalizing this, firm markdowns are wider, and employment and output are lower at all firms. With firm heterogeneity, progressivity also misallocates labor across firms: progressive taxes make labor relatively more expensive at higher wage, higher productivity firms. Hence monopsony delivers a novel channel through which progressive taxes themselves lead to *inefficiency*, despite being potentially beneficial from a *redistributive* standpoint.[30] Of course, here we abstract from the *insurance* benefits of progressive taxes. Ongoing work adds idiosyncratic risk in a Bewley economy to understanding the extent to which this new inefficiency may off-set insurance benefits (Berger, Herkenhoff, and Mongey, 2023).

Implementation. We recalibrate shifters $\{\xi_h, \widetilde{\varphi}_h, \overline{Z}, \overline{\varphi}\}$ to exactly match the same moments in Table 4, but now in terms of pre-tax wages. Rather than take a stand on whether the subsidy and tax system should be balanced, we prioritize matching the shape of the tax system (Figure 10A). Under $\underline{w} = 0$ and $(\tau, \lambda) = (0.181, 1.746)$ the tax system delivers a small surplus of $g = 0.88\%$ of output (i.e. $G = Taxes - Subsidies = gY$), which now enters the resource constraint. We fix $g = 0.88$, and at

[28]We use the tax schedule for single households. This varies by number of children. We average across the distribution of number of children. Data are from Congressional Research Service report *"The Earned Income Tax Credit (EITC): How It Works and Who Receives It."* (January, 2021)

[29]See Appendix A.I for the derivation.

[30]We clarify these theoretical points in Berger, Herkenhoff, Mongey, and Mousavi (2024).

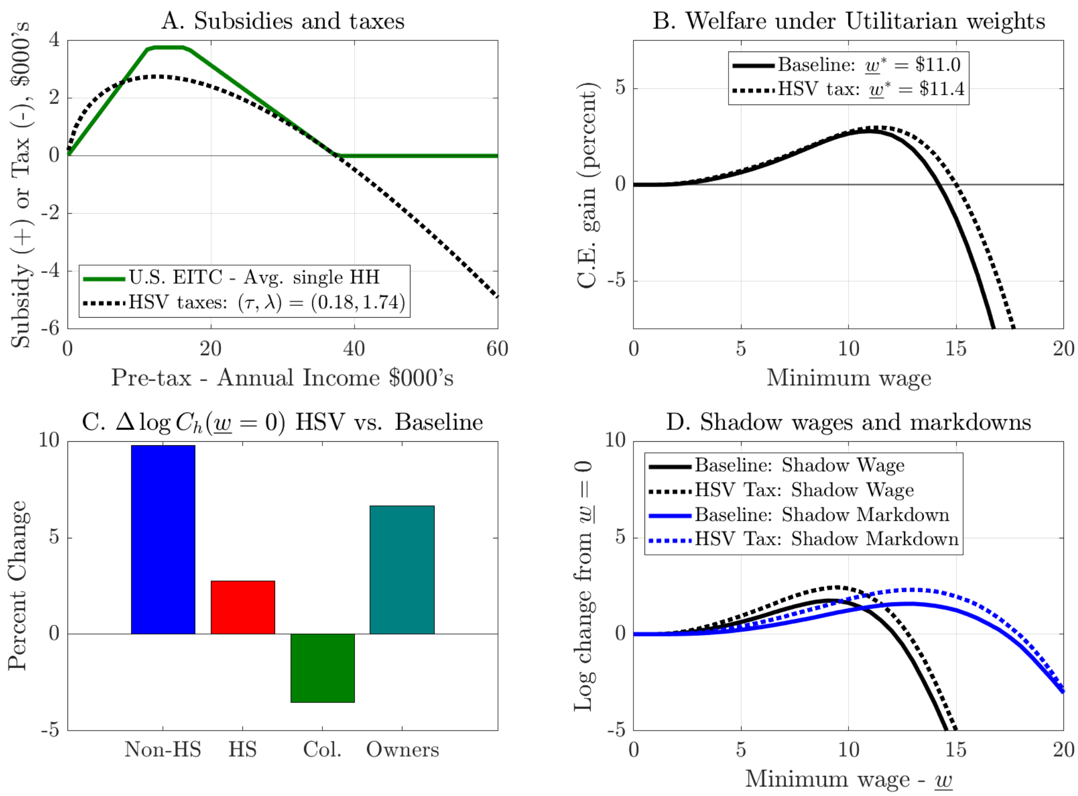

Figure 10: Efficiency of the minimum wage under progressive taxes

each $\underline{w}$ solve for the λ that clears the government budget constraint.[31]

Optimal minimum wage with taxes and transfers. Intuition would suggest that with greater redistribution, the optimal minimum wage should fall towards that of the homogeneous worker economy. However, this is not the case. Figure 10B plots Utilitarian welfare gains with and without the subsidy and tax system. The optimal minimum wage and welfare gains barely change, but both slightly *increase*.

As discussed above, progressive taxes exacerbate monopsony power. With more monopsony power, business owner profits increase, which is at odds with what a Utilitarian planner would like to achieve. Figure 10C shows how consumption changes between the baseline and HSV economies. Consistent with the redistributive role of the tax system, non-college households consume more, and college workers consume less. However, business owner consumption *rises* as progressive taxes distort wage setting power. Overall, the redistributive force

[31]This follows the approach in Boerma and Karabarbounis (2021) to analyzing alternative tax policies.

of the minimum wage—which is to transfer resources from business owners to non-business owners—remains in tact, but is muted.

Facing effectively less elastic labor supply in response to pre-tax wages (equation 15), markdowns are wider, and hence the minimum wage has more scope to improve welfare. Under HSV taxes, Figure 10D shows that shadow wages and shadow markdowns improve by more and peak at higher minimum wages. This puts a small amount of upward pressure on the Utilitarian optimal minimum wage.

Welfare gains from minimum wages vs. taxes. The welfare gains from minimum wages are small relative to the efficient allocation. Would an optimal HSV tax system – denoted τ^* and λ^* – deliver more of the potential redistributive and/or efficiency gains? No. Holding government spending-to-output constant and setting $\underline{w} = 0$, we find that the optimal degree of progressivity and subsidy/tax cutoff are $\tau^* = 0.29$ and $\lambda^* = 2.39$. This is more progressive than the empirical baseline of $\tau = 0.18$ and yields a larger threshold for receipt of a net subsidy than the empirical baseline of $\lambda = 1.74$. However, the overall Utilitarian welfare gain from (τ^*, λ^*) relative to the baseline (τ, λ) is 1.83%, which remains dwarfed by the 30% gains from the efficient allocation.

The optimal policy yields efficiency losses of approximately 3%, and redistributive gains of approximately 4%. The efficiency losses from progressive taxes are unsurprising. But what limits scope for redistribution via progressive taxation? Equation (15) shows that greater progressivity yields more labor market power for business owners. They charge greater markdowns and consume more. At progressivity rates beyond $\tau = 0.70$, widening markdowns yield redistributive *losses*.

These results are subject to several caveats: (1) optimal progressivity depends critically on welfare weights, and the focus of our paper is on efficiency, not redistribution, (2) the Negishi weights that rationalize current tax policy are far from Utilitarian, and (3) we abstract from important insurance motives present in most optimal tax exercises, e.g. Heathcote, Storesletten, and Violante (2017). We provide more details in Appendix A.I.

Implications for inequality. Our final exercise explores the effects of minimum wages on standard metrics for inequality – the college wage premium and the variance of log wages – and asks whether these redistributive metrics are useful for guiding policy.

Figure 11 shows that the minimum wage has powerful effects on both margins. Panel A plots the pre-tax (solid) and post-tax (dashed) premium of college workers' average wage relative to (a) non-high school workers (upper lines, red), and (b) all non-college workers (lower lines, black). Raising the minimum wage to $20 reduces the post-tax premium relative to non-high school workers by 41 log points and all non-college workers by 22 log points. Panel B shows that a $20 minimum wage also reduces after-tax wage inequality by more than 15 log points.

Are wage premia and inequality metrics useful for guiding policy makers interested in redistribution? We argue no. Both wage premia and wage dispersion are monotonically declining in the minimum wage, despite the single-peaked welfare of each household type. Take for instance a planner that values only redistribution toward the lowest income households in the economy: non-highschool graduates. Panel A says that a policy that minimizes the gap between college and non-highschool wages would yield a minimum wage in excess of $20. Panel C says that such a policy prescription would be at odds with even the most extreme preferences for redistribution towards non-highschool workers. We plot welfare for the lowest and highest earning non-HS worker households. A planner that places all social welfare weight on the lowest (highest) non-highschool earner would set a minimum wage of $8.50 ($18.00). No non-highschool household would choose a minimum wage of $20, despite it narrowing the inequality between these households and college households. We conclude that standard metrics for inequality have little normative value, regardless of the objective of the planner.

7 Discussion

Our results are that the efficiency gains from minimum wages are low, and gains that exist under Utilitarian social welfare weights are almost entirely driven by redistribution. We provide further support for these results in three ways. First, we show that low efficiency gains are not due to the model insufficiently capturing channels for improved efficiency pointed to by the empirical literature. We replicate leading empirical studies on the spillover, reallocation and employment

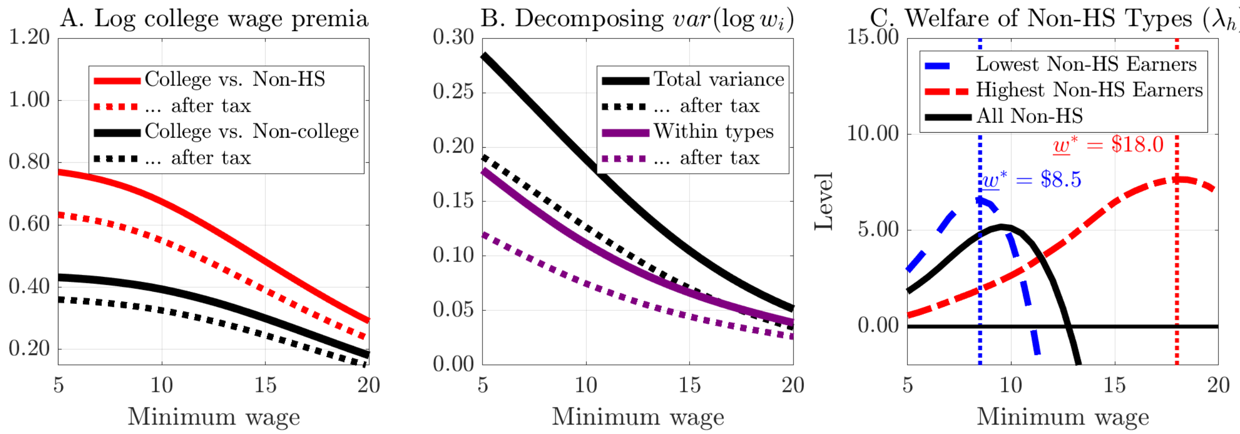

Figure 11: Minimum wages and commonly used empirical proxies for welfare

effects of minimum wages, and how these interact with market structure. Second, we provide robustness with respect to our model parameters and calibration strategy. Third, we reason that incorporating missing features would push toward lower efficiency gains.

Validation. There are three channels through which minimum wages may improve efficiency: (i) direct narrowing of markdowns, (ii) wage spillovers which undo distortions at unconstrained firms, and (iii) reallocation to more productive firms. We validate our model's responses of each of these channels to minimum wages by replicating four recent studies in Appendix O.B. We first assess our model's direct effects of minimum wages on employment and wages by replicating a recent study of small and large minimum wage hikes in Seattle (Jardim, Long, Plotnick, Van Inwegen, Vigdor, and Wething (2022)). We then study how the model's direct effects vary by market concentration by replicating Azar et. al. (2023). They find employment gains in concentrated markets, a feature that is only reproduceable by models such as our with variable markdowns (with common markdowns, gains are independent of concentration). Engbom and Moser (2022) use detailed hours and earnings data from Brazil to measure spillovers, avoiding measurement error issues that plague studies in the U.S. We generate quantitatively and qualitatively similar spillover patterns.[32] Lastly, Dustmann et. al. (2022) study reallocation of workers between firms in Germany. Our model replicates the

[32]In BHM we quantitatively replicated Staiger, Spetz, and Phibbs (2010), which documented how competing hospitals raised nurse's wages following the imposition of a wage floor at Veteran's Affairs hospitals in 1991.

reallocation of workers from smaller to larger firms as minimum wages rise. In summary, the model successfully replicates and gives a natural interpretation to key reduced form results from the empirical literature on minimum wages.

Robustness exercises. Appendix O.C provides details of the following robustness exercises. First, for a wide range of Frisch elasticities $\varphi \in [0.3, 0.9]$, the efficiency maximizing minimum wage, and the resulting gains, are effectively unchanged. Second, we find very little heterogeneity in efficiency gains or efficiency maximizing minimum wages across regions. We calibrate to low income states, high income states, and Mississippi, and in all cases the Utilitarian welfare gains lie between 2.70% and 2.80%, and the aggregate efficiency gains lie between 0.05% and 0.11%. Third, the inclusion of capital accommodates an exercise in which we assume the capital each firm allocates to each worker type is fixed as the minimum wage increases. In this case firms will want to shutdown as capital expenses become a fixed overhead cost, and hence we solve for the equilibrium with an endogenous amount of exit. In this exercise, the efficiency maximizing minimum wage falls by only 27 cents. Fixed capital steepens decreasing returns to labor, narrowing Region II, and reducing the scope of $\underline{w}$ to expand employment. Fourth, we consider lower degrees of substitutability across labor types. We re-calibrate α (recall, equation 9) to deliver an elasticity of 2.9 (Acemoglu and Autor, 2011). The efficiency maximizing minimum wage falls by about 50 cents. Fifth, we argue the effects of lower capital-labor substitutability (e.g. Oberfield and Raval, 2021) can be bound by our fixed capital exercise, yielding an extreme elasticity of substitution of zero.

Finally, we reduce the amount of worker heterogeneity by calibrating a model with only a single non-highschool, highschool and college household (i.e. four household types in total). First, as expected, the efficiency maximizing minimum wage is barely changed, consistent with our earlier results that the homogeneous worker and heterogeneous worker economies deliver the same answer with respect to the efficiency maximizing minimum wage, which is the focus of this paper. With four types the efficiency maximizing minimum wage is \$7.18, compared to \$7.35 in our baseline twelve type calibration. Second, the Utilitarian optimal minimum wage that maximizes overall welfare, inclusive of redistribution and efficiency, is barely changed: \$10.53, compared to \$11.00 in our baseline twelve type calibration. We conclude that our results are robust to a simplified view of heterogeneity in the economy, and leave it to future work to understand whether much richer heterogeneity changes these results.

Discussion of missing features. Our model necessarily omits a number of features: pass-through to prices, automation, a non-unitary elasticity of substitution between capital and labor, incomplete markets and borrowing constraints, and inefficient rationing. We discuss each feature and argue that including each will likely lead to even smaller efficiency and redistributive welfare gains.

First, quantitative models of product market competition imply firms facing less competition charge the widest markups (e.g. Edmond, Midrigan, and Xu, 2023).[33] Minimum wages first raise marginal costs at small firms which delivers more product market power to large firms, compounding distortions.[34] Second, automation and higher substitutability between capital and labor will fossilize any short-term rationing that occurs, which will again reduce welfare gains.[35] Third, incomplete markets and borrowing constraints would further cut into any benefits from raising the minimum wage. This would particularly bite in a life-cycle model with human capital accumulation or in a model with uninsurable unemployment risk. Fourth, we do not consider inefficient rationing, since within households workers are homogeneous.[36] Inefficient rationing would further limit efficiency gains and compound efficiency losses.

Finally, we note that our model does not allow for work below the minimum wage, while in the CPS some wages below the minimum wage are observed. It is unclear whether these are wages from measurement error or informal work—

[33]Our framework is fungible enough to include imperfect competition in the production market. Our benchmark model incorporates a decreasing marginal revenue product of labor through decreasing returns in production, but could be replaced by downward sloping demand under monopolistic competition.

[34]This presents a cynical view of the full-page newspaper advertisements purchased by Amazon in 2021 encouraging U.S. Congress to pass a Federal \$15 minimum wage law. See coverage of the anti-competitive implications for lower wage competitors here: *Amazon's Push for a \$15 Minimum Wage is a New Weapon in Company's Battle Against Walmart* (Business Insider, February 24, 2021.

[35]Harasztosi and Lindner (2019) documents firms substitute away from labor and toward capital, increasing purchases of computers and other capital goods.

[36]The canonical example being a \$15/hour minimum wage job that ends up going to a worker that would work for \$14/hour while a worker that would work for \$10/hour remains unemployed.

which would be a pressing matter if extending our work to developing countries.

8 Conclusion

This paper provides a theoretical framework for studying the effect of minimum wages on welfare and the allocation of employment across firms in the economy. The framework has three key features. First, each market features strategic interaction between firms, which we have shown to be important for (i) quantifying the reallocation effects of minimum wage policies, (ii) interpreting empirical evidence documenting such reallocation, and (iii) interpreting empirical evidence on spillovers of minimum wages. Second, workers are of heterogeneous types, which allows us to decompose the heterogeneous impacts of minimum wages on employment, labor and capital income, and account for general equilibrium wealth effects. Third, we provide a parsimonious nesting of this market model into a general equilibrium economy and show how the economy aggregates, allowing for a succinct representation of the efficiency improvements and costs of minimum wages via *shadow markdown* $\widetilde{\mu}$, and misallocation ω. When calibrated to US data, the model is consistent with a wide body of empirical research on the effects of minimum wage changes.

In such an economy we find that an optimal minimum wage exists. Quantitatively, we find that the efficiency maximizing minimum wage is less than $8 per hour, but that higher minimum wages can be justified through redistribution, even under a redistributive tax and transfer system.

References

AARONSON, D., E. FRENCH, I. SORKIN, AND T. TO (2018): "Industry Dynamics and the Minimum Wage: A Putty-clay Approach," *International Economic Review*, 59(1), 51–84.

ACEMOGLU, D., AND D. AUTOR (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter, vol. 4B, pp. 1043–1171. North-Holland.

AHLFELDT, G. M., D. ROTH, AND T. SEIDEL (2022): "Optimal Minimum Wages," CEPR Discussion Paper DP16913, CEPR.

AZAR, J., E. HUET-VAUGHN, I. MARINESCU, B. TASKA, AND T. VON WACHTER (2023): "Minimum Wage Employment Effects and Labour Market Concentration," *The Review of Economic Studies (forthcoming)*.

BARKAI, S. (2020): "Declining Labor and Capital Shares," *The Journal of Finance*, 75(5), 2421–2463.

BENABOU, R. (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 70(2), 481–517.

BERGER, D. W., K. F. HERKENHOFF, AND S. MONGEY (2022): "Labor Market Power," *American Economic Review*, 112(4).

——— (2023): "Labor Market Power, Tax Progressivity and Inequality," Slides available at http://simonmongey.com, Working Paper.

——— (2024): "Supplement to Minimum Wages, Efficiency and Welfare," .

BERGER, D. W., K. F. HERKENHOFF, S. MONGEY, AND N. MOUSAVI (2024): "Monopsony amplifies distortions from progressive taxes," Working Paper 31972, National Bureau of Economic Research.

BHASKAR, V., AND T. TO (1999): "Minimum Wages for Ronald McDonald Monopsonies: A Theory of Monopsonistic Competition," *The Economic Journal*, 109(455), 190–203.

BLOOM, N., F. GUVENEN, B. S. SMITH, J. SONG, AND T. VON WACHTER (2018): "Inequality and the disappearing large firm wage premium," in *American Economic Review Papers & Proceedings*, vol. 108, pp. 317–22.

BOERMA, J., AND L. KARABARBOUNIS (2021): "Inferring Inequality With Home Production," *Econometrica*, 89(5), 2517–2556.

BURDETT, K., AND D. MORTENSEN (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2), 257–73.

CAHUC, P., AND G. LAROQUE (2014): "Optimal Taxation and Monopsonistic Labor Market: Does Monopsony Justify the Minimum Wage?," *Journal of Public Economic Theory*, 16(2), 259–273.

CAHUC, P., AND A. ZYLBERBERG (2004): *Labor Economics*. MIT Press, Cambridge, MA.

CENGIZ, D., A. DUBE, A. LINDNER, AND B. ZIPPERER (2019): "The Effect of Minimum Wages on Low-Wage Jobs*," *The Quarterly Journal of Economics*, 134(3), 1405–1454.

CHAN, M., E. MATTANA, S. SALGADO, AND M. XU (2023): "Wage Setting and Passthrough: The Role of Market Power, Production Technology and Adjustment Costs," Working paper, Queen's University.

CHING, A. T., F. HAYASHI, AND H. WANG (2015): "Quantifying the Impacts of Limited Supply: The Case of Nursing Homes," *International Economic Review*, 56(4), 1291–1322.

DÁVILA, E., AND A. SCHAAB (2022): "Welfare Assessments with Heterogeneous Individuals," Discussion Paper 30961, National Bureau of Economic Research.

DE PALMA, A., N. PICARD, AND P. WADDELL (2007): "Discrete Choice Models with Capacity Constraints: An Empirical Analysis of the Housing Market of the Greater Paris Region," *Journal of Urban Economics*, 62(2), 204–230.

DECKER, R. A., J. HALTIWANGER, R. S. JARMIN, AND J. MIRANDA (2020): "Changing Business Dynamism and Productivity: Shocks Versus Responsiveness," *American Economic Review*, 110(12), 3952–90.

DERENONCOURT, E., AND C. MONTIALOUX (2021): "Minimum Wages and Racial Inequality," *The Quarterly Journal of Economics*, 136(1), 169–228.

DUSTMANN, C., A. LINDNER, U. SCHÖNBERG, M. UMKEHRER, AND P. VOM BERGE (2022): "Reallocation Effects of the Minimum Wage," *The Quarterly Journal of Economics*, 137(1), 267–328.

EDMOND, C., V. MIDRIGAN, AND D. Y. XU (2023): "How Costly Are Markups?," *Journal of Political Economy*, 131(7).

ENGBOM, N., AND C. MOSER (2022): "Earnings inequality and the minimum wage: Evidence from Brazil," *American Economic Review*, 112(12), 3803–47.

FLINN, C. (2006): "Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates," *Econometrica*, 74(4), 1013–1062.

——— (2010): *The Minimum Wage and Labor Market Outcomes*. MIT Press, Cambridge, MA.

FLINN, C., AND J. MULLINS (2021): "Firms' Choices of Wage-Setting Protocols," Discussion paper, New York University.

FLODEN, M. (2001): "The Effectiveness of Government Debt and Transfers as Insurance," *Journal of Monetary Economics*, 48(1), 81–108.

GOLOSOV, M., M. GRABER, M. MOGSTAD, AND D. NOVGORODSKY (2021): "How Americans Respond to Idiosyncratic and Exogenous Changes in Household Wealth and Unearned Income," Working Paper 29000, National Bureau of Economic Research.

HAANWICKEL, D. (2023): "Supply, Demand, Institutions, and Firms: A Theory of Labor Market Sorting and the Wage Distribution," Working Paper 31318.

Harasztosi, P., and A. Lindner (2019): "Who Pays for the Minimum Wage?," *American Economic Review*, 109(8), 2693–2727.

Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal Tax Progressivity: An Analytical Framework," *The Quarterly Journal of Economics*, 132(4), 1693–1754.

Hurst, E., P. Kehoe, E. Pastorino, and T. Winberry (2022): "The Distributional Impact of the Minimum Wage in the Short and Long Run," Working Paper 30294, National Bureau of Economic Research.

Janez, A., and L. Delgado-Prieto (2023): "Monopsony Power and Firm Organization," Working paper, Universidad Carlos III de Madrid.

Jardim, E., M. Long, R. Plotnick, E. Van Inwegen, J. Vigdor, and H. Wething (2022): "Minimum-Wage Increases and Low-Wage Employment: Evidence from Seattle," *American Economic Journal: Economic Policy*, 14(2), 264–314.

Katz, L. F., and K. M. Murphy (1992): "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107(1), 35–78.

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

MaCurdy, T. E. (1981): "An Empirical Model of Labor Supply in a Life-cycle Setting," *Journal of Political Economy*, 89(6), 1059–1085.

Marhsall, S. (2023): "Monopsony in Sub-Saharan Africa: The Effect of Small Firms, Self-Employment, and Migration," Working paper, University of Warwick.

Molho, I. (1995): "Price Ceilings and Imperfect Competition," *Scottish Journal of Political Economy*, 42(4), 392–408.

Oberfield, E., and D. Raval (2021): "Micro Data and Macro Technology," *Econometrica*, 89(2), 703–732.

Reynolds, S. S., and D. Rietzke (2018): "Price Caps, Oligopoly, and Entry," *Economic Theory*, 66, 707–745.

Robinson, J. (1933): *The Economics of Imperfect Competition*. Palgrave Macmillan, London.

Sorkin, I. (2015): "Are There Long-run Effects of the Minimum Wage?," *Review of Economic Dynamics*, 18(2), 306–333.

Staiger, D. O., J. Spetz, and C. S. Phibbs (2010): "Is There Monopsony in the Labor Market? Evidence from a Natural Experiment," *Journal of Labor Economics*, 28(2), 211–236.

VOGEL, J. (2022): "The Race Between Education, Technology and Institutions," Working Paper 30311, National Bureau of Economic Research.

YEH, C., C. MACALUSO, AND B. HERSHBEIN (2022): "Monopsony in the US Labor Market," *American Economic Review*, 112(7), 2099–2138.