Minimum Wages, Efficiency and Welfare*

David Berger Kyle Herkenhoff Simon Mongey

April 24, 2024

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 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. We thank Anmol Bhandari, Jeffrey Clemens, Elora Derenoncourt, Arindrajit Dube, Erik Hurst, Ioana Marinescu, Loukas Karabarbounis, Patrick Kline and Atilla Lindner for helpful conversations. 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 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 (i) an *Online Appendix*, published by this journal, and (ii) *Supplemental Appendix* published as a separate working paper found on the journal's website (Berger, Herkenhoff, and Mongey, 2024). We refer to these as Appendix O and Appendix S, 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 maximum amount of labor the firm is willing to hire, hence $n_{ij} \leq \overline{n}_{ij}$. We call this a *rationing constraint*.

[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.

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,

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

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.

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

[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).

in equation (6) are now written in terms of the shadow wage-bill share (proof in Appendix O.D):

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

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:

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

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 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

[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.

[10] In Region II, the marginal cost curve is different from the benchmark economy. The new

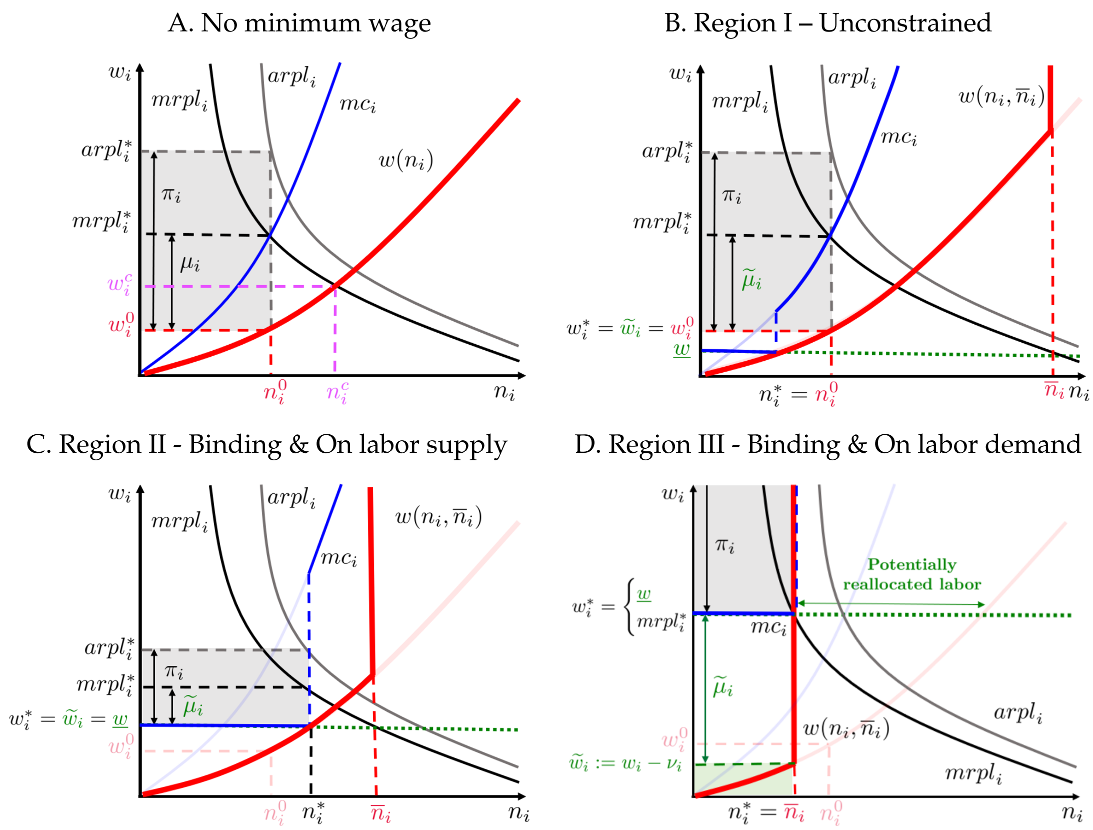


Figure 1: Increase in $\underline{w}$ - Partial equilibrium

Notes: The dashed green line corresponds to the minimum wage $\underline{w}$. The red line 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 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 $n_i \in (0, \overline{n}_i]$.

$(p_i = 1)$, 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 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*

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 $\underline{w}$.

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 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 by the first green 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

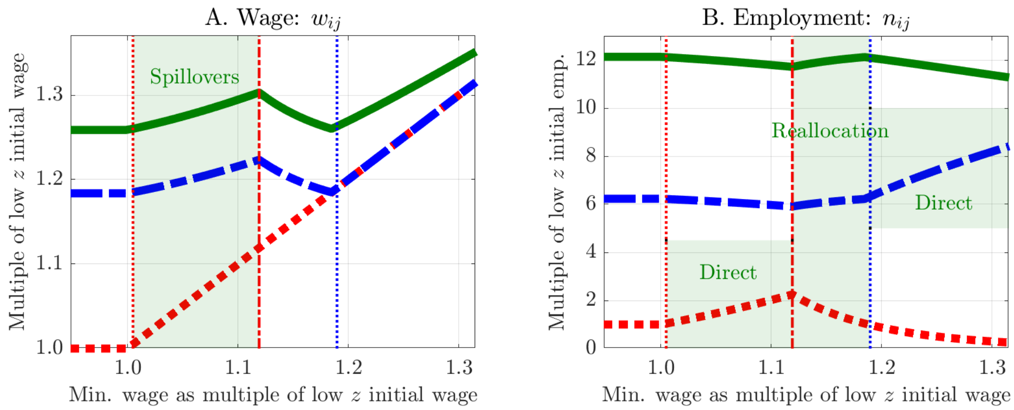


Figure 2: Increase in $\underline{w}$ - Firm outcomes in market equilibrium

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 $z^{low} = 1.97$ (red, short dash), $z^{med} = 4.04$ (blue, long-dash), $z^{high} = 6.42$ (green, solid). National W and N are held fixed at value corresponding to $\underline{w} = 0$.

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 red 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 green 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 red 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 S.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 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 red 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 red firm is in Region III and its competi-

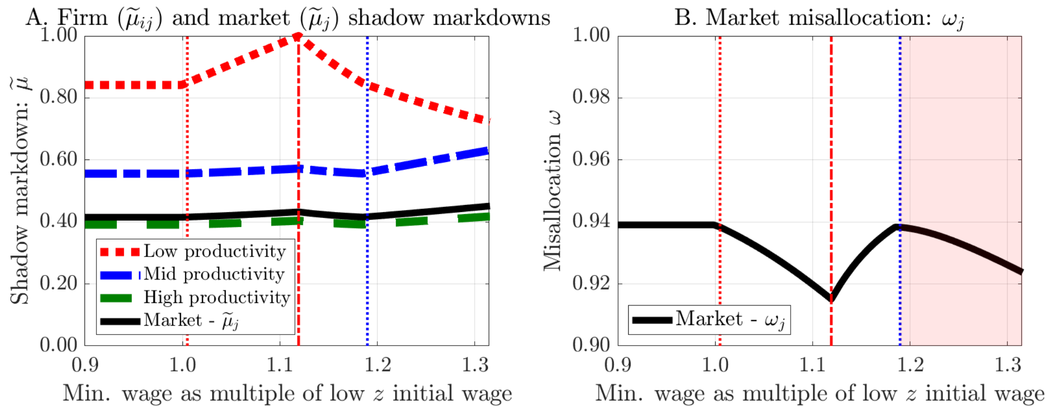


Figure 3: Increase in $\underline{w}$ - Market outcomes - Shadow markdown and misallocation

Notes: The model economy is identical to Figure 2. The productivities are given by $z^{low} = 1.97$ (red, short dash), $z^{med} = 4.04$ (blue, long-dash), $z^{high} = 6.42$ (green, solid). Panel *A* plots the market markdown $\widetilde{\mu}_j$ (black solid). Panel *B* plots the market misallocation $\widetilde{\omega}_j$ (black solid). 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).

tors are unconstrained. However, it worsens once the medium productivity firm starts paying the minimum wage (shaded in red). The green firm 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 effi-

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

ciency 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

[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 S.E.

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 $M_j > 1$. 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

[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. $IQR = \log(z(p75)/z(p25))$) 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.

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 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.

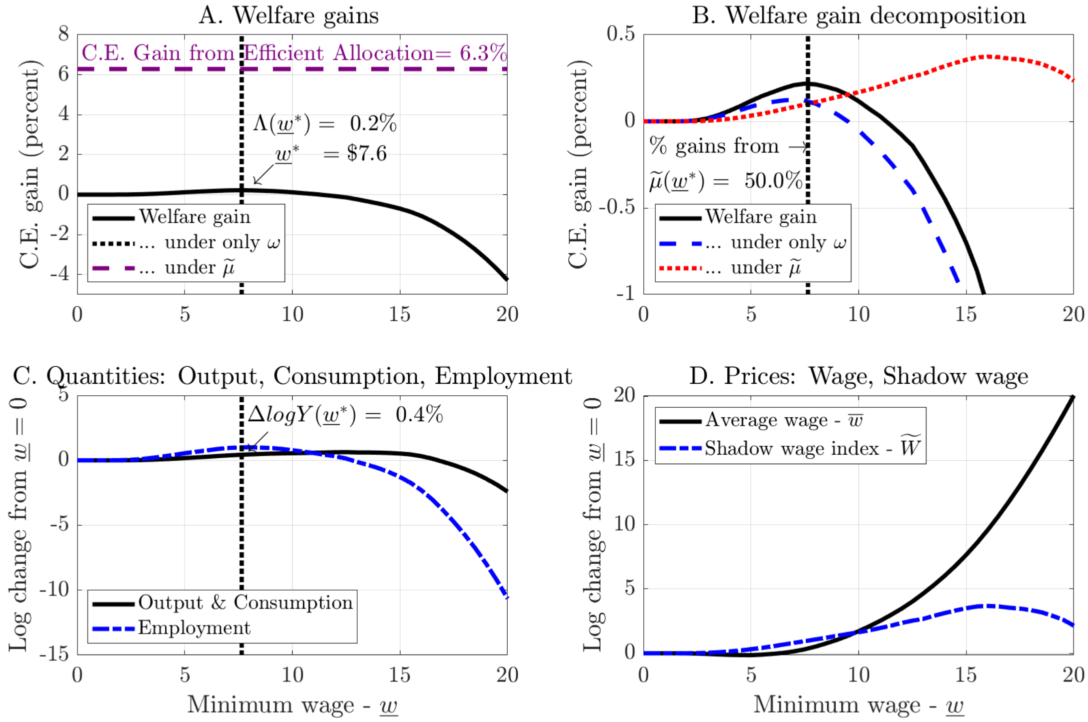


Figure 4: Minimum wages and welfare

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 allocational efficiency ω_k only. The short-dash 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).

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

[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).

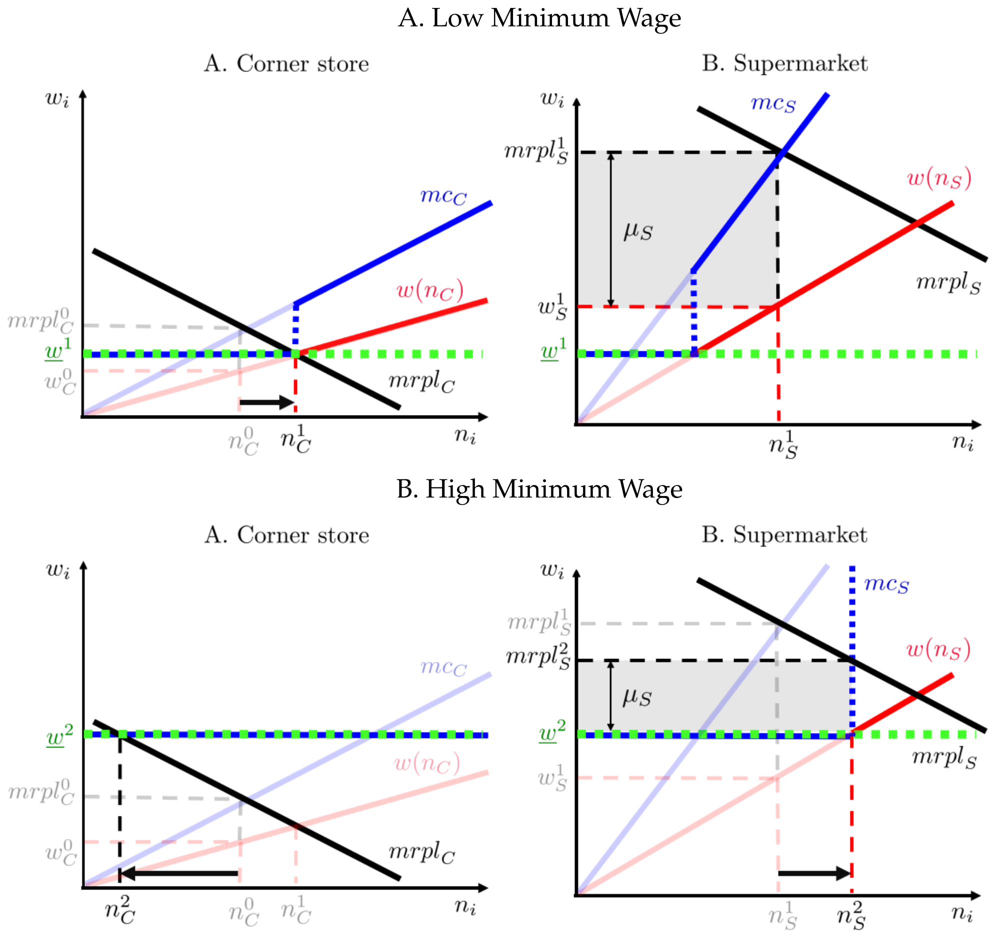


Figure 5: Productivity heterogeneity limits efficiency gain from minimum wage

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 ($\eta = 10$).

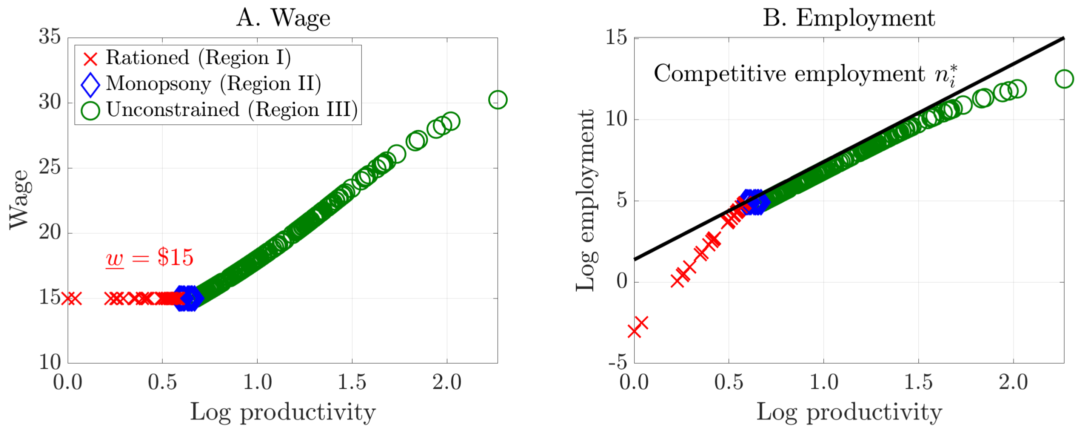


Figure 6: Small efficiency gains from minimum wage in a 200 firm market

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. Red $\times$'s are Region I firms, blue diamonds are Region II firms, and green circles are Region III firms. The black line 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 $\widetilde{\mu}_{ij} = 1$: $w_{ij} = \widetilde{\alpha}\widetilde{z}_{ij} n_{ij}^{\widetilde{\alpha}-1}$.

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 blue 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.

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. Red crosses 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-

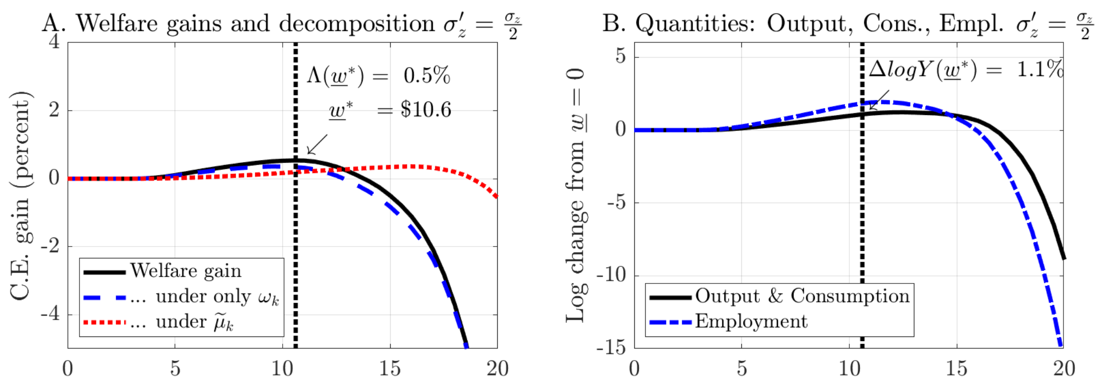


Figure 7: Minimum wages and welfare - Half dispersion in productivity

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 allocational efficiency ω_k only. The short-dash 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).

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 in the green 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 Supplemental Appendix I, 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.

	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

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.

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

[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 Table A1.

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.

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 and C.5 for discussion). The labor-labor elasticity of substitution between types h and h' is

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

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

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

[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

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, $n_{ijht} \leq \overline{n}_{ijht}$.

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}:

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

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:

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

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 $\overline{\varphi}$: $\widetilde{\varphi}_h = (\overline{\varphi}_h/\overline{\varphi})\, \pi_h^{1+\varphi}$

[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.

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

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

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 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.

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.

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]

[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.

Model inversion. We first take population shares π_h from the CPS. Parameters that are heterogeneous across households are relative shifters in productivity and labor supply disutility $\{\zeta_h, \widetilde{\varphi}_h\}_{h=1}^H$, 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 $\zeta_h = \widetilde{\varphi}_h = 1$ 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 $\{\zeta_h\}_{h=1}^H$ 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 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).

[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.

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.

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

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

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 ($\mu_{ijh} = 1, \ \forall ijh$). 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

[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.

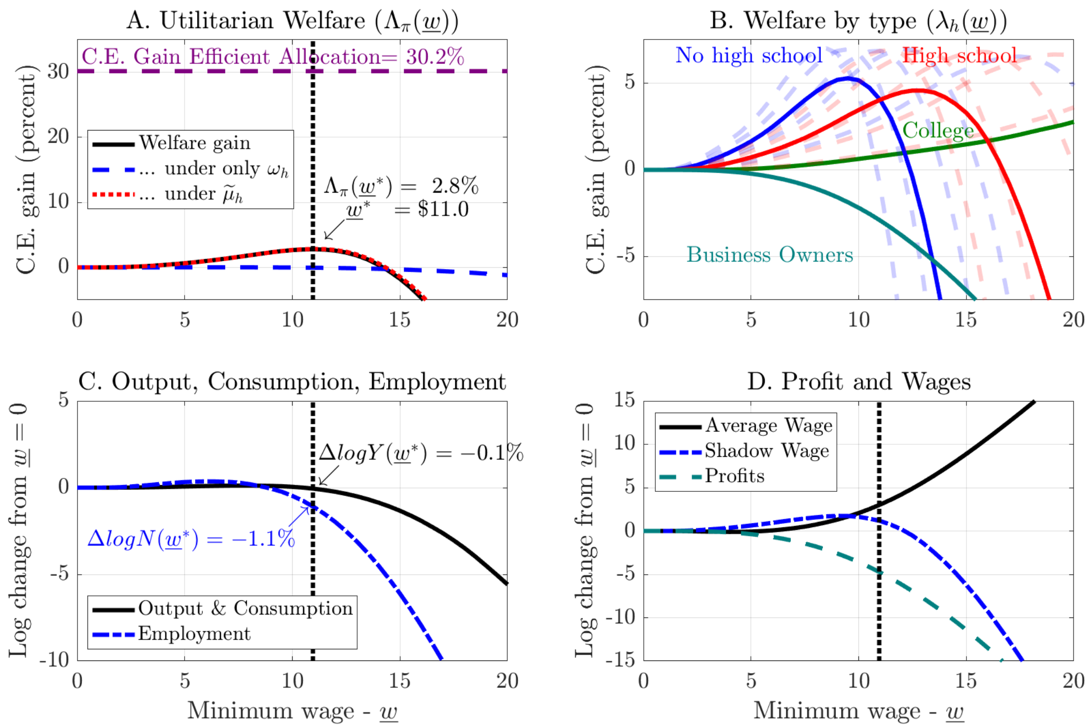


Figure 8: Minimum wages and welfare

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 $\Lambda_\pi(\underline{w})$ (black line) attributable to markdowns (red dashed line) and misallocation (blue dashed line). The efficient allocation welfare gains are denoted by the horizontal purple line 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 (blue), high school workers (red), college workers (green), and business owners (teal) **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 black dashed vertical line. **Panel D.** Plots average wages (black solid line), the average wage index across worker types (blue dash-dot line), and business profits (teal dashed line).

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:

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

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 $\frac{RE}{TOT}$.

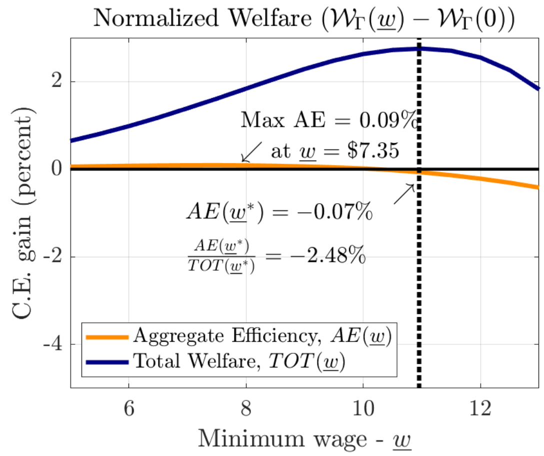


Figure 9: Minimum wages and welfare

Notes: This figureplots 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.

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.

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

[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 S.I for the derivation.

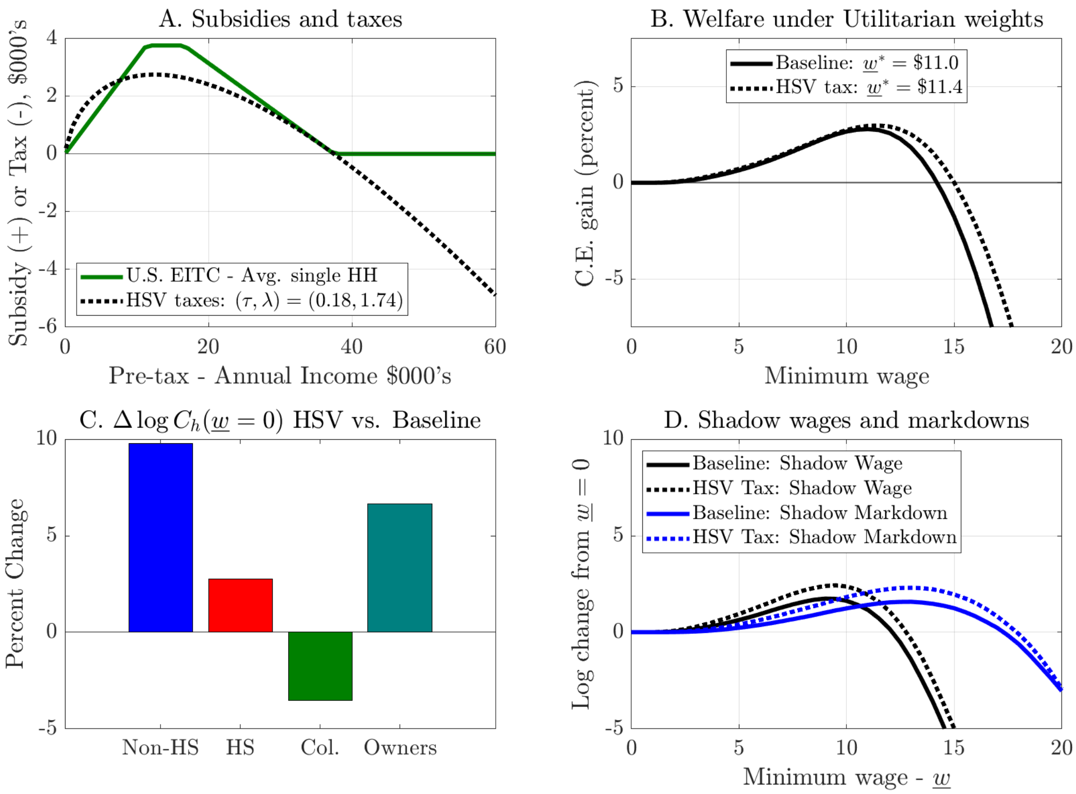


Figure 10: Efficiency of the minimum wage under progressive taxes

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 $(\tau, \lambda) = (0.181, 1.746)$

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

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 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. The redistributive force of the minimum wage—which is to transfer resources from business owners to non-business owners—remains in tact.

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 for improve welfare. Figure 10D shows that shadow wages and shadow markdowns improve by more and under higher minimum wages with HSV taxes. 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

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

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 and redistributive gains of approximately -3% and 4%, respectively. 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. Widening markdowns yield redistributive *losses* at progressivity rates beyond $\tau = 0.7$.

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 Supplemental Appendix 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 (red solid), and (b) all non-college workers (black solid). 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

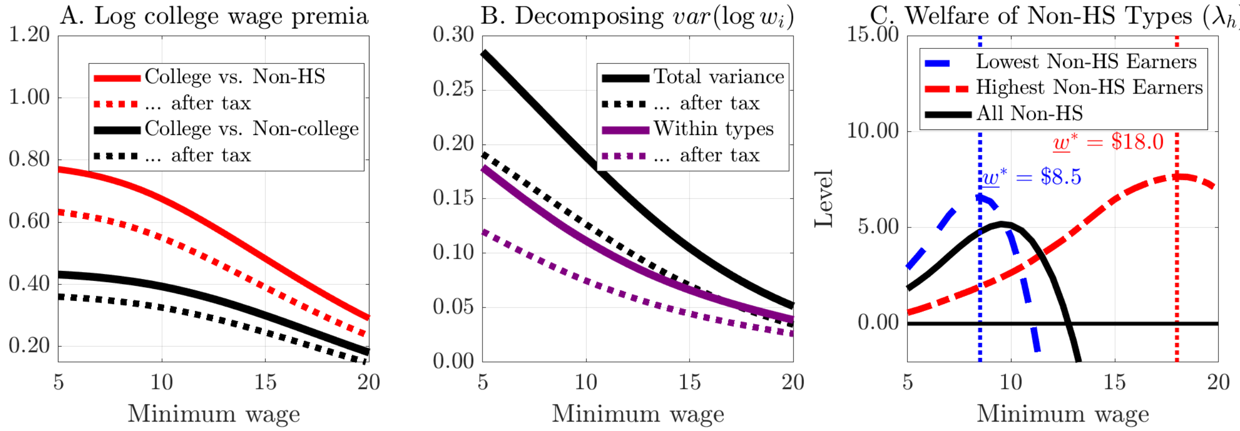


Figure 11: Minimum wages and commonly used empirical proxies for welfare

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 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 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. Appendices 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, nar-

[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.

rowing 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

[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.

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—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

[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.

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.

AUTOR, D. H., A. MANNING, AND C. L. SMITH (2016): "The Contribution of the Minimum Wage to US Wage Inequality over Three Decades: A Reassessment," *American Economic Journal: Applied Economics*, 8(1), 58–99.

AZAR, J., 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): "Supplemental Appendix 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.

CARD, D., AND T. LEMIEUX (2001): “Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-based Analysis,” *The Quarterly Journal of Economics*, 116(2), 705–746.

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.

CLEMENS, J., AND M. R. STRAIN (2021): “The Heterogeneous Effects of Large and Small Minimum Wage Changes: Evidence over the Short and Medium Run Using a Pre-Analysis Plan,” Working Paper 29264, National Bureau of Economic Research.

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.

GECHERT, S., T. HAVRANEK, Z. IRSOVA, AND D. KOLCUNOVA (2022): "Measuring Capital-labor Substitution: The Importance of Method Choices and Publication Bias," *Review of Economic Dynamics*, 45, 55–82.

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. VIOLANTE (2020): "How Should Tax Progressivity Respond to Rising Income Inequality?," *Journal of the European Economic Association*, 18(6).

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.

KARABARBOUNIS, L., AND B. NEIMAN (2014): "The Global Decline of the Labor Share," *The Quarterly Journal of Economics*, 129(1), 61–103.

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.

NEUMARK, D., AND P. SHIRLEY (2022): "Myth or Measurement: What Does the New Minimum Wage Research Say About Minimum Wages and Job Loss in the United States?," *Industrial Relations: A Journal of Economy and Society*, 61(4), 384–417.

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.

Online Appendix

Section A provides model parameters and moments as well as the wage distribution in the model and data. Section B provides details of the ***Validation*** exercises described in Section 7. Section C provides ***Robustness*** exercises described in Section 7. Section D contains ***Proofs*** for a simplified monopsony and oligopsony economy that are referred to in Section 3, and an even simpler pedagogical example. This is the Homogeneous worker economy. The *Supplemental Appendix* follows the *Online Appendix*, and provides (i) details on the calibration of φ, (i) additional figures and tables, (ii) derivations of the equilibrium conditions for the Heterogeneous worker economy, and any other equations in the main text, (iii) algorithm for solving the economy.

A Additional calibration details and fit

Table A1 provides the full set of parameters for all 12 types of households. Table A2 reports the detailed moments. Figure A1 plots the wage PDF for model v. data in the 12 type economy.

Parameters		**NHS**					**HS**					**C**	**O**
Wage quintile		1	2	3	4	5	1	2	3	4	5	All	All
Relative population (%)	$\pi_h / \sum \pi_h$	2.6	2.6	2.6	2.6	2.6	10.7	10.7	10.7	10.7	10.7	26.1	7.0
Relative disutility labor supply	$\overline{\varphi}_h^{-\varphi} / \overline{\varphi}_C^{-\varphi}$	39.13	6.17	3.70	2.78	1.98	4.80	2.07	1.68	1.44	1.05	1.00	0.53
Relative productivity	ξ_h	0.14	0.18	0.23	0.27	0.42	0.27	0.33	0.42	0.55	0.87	1.00	0.89
Fraction of capital (%)	κ_h	0.00	0.00	0.01	0.01	0.07	0.04	0.13	0.16	0.27	1.04	4.30	93.96

Table A1: Detailed Parameters

	Model											
Targets (* Means Model=Data)	**NHS**					**HS**					**C**	**O**
Wage quintile	1	2	3	4	5	1	2	3	4	5	All	All
Population shares* (CPS, %)	2.64	2.64	2.64	2.64	2.64	10.74	10.74	10.74	10.74	10.74	26.08	7.00
Share of agg. labor income* (CPS and SCF, %)	0.04	0.26	0.50	0.74	1.46	1.64	4.26	6.28	8.92	17.36	46.16	12.39
Ave. earnings per hour*, (CPS, C=1)	0.29	0.32	0.37	0.41	0.60	0.39	0.44	0.54	0.66	0.95	1.00	1.00
Capital income to labor income* (SCF)	0.00	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.03	0.05	4.00

Table A2: Detailed Moments

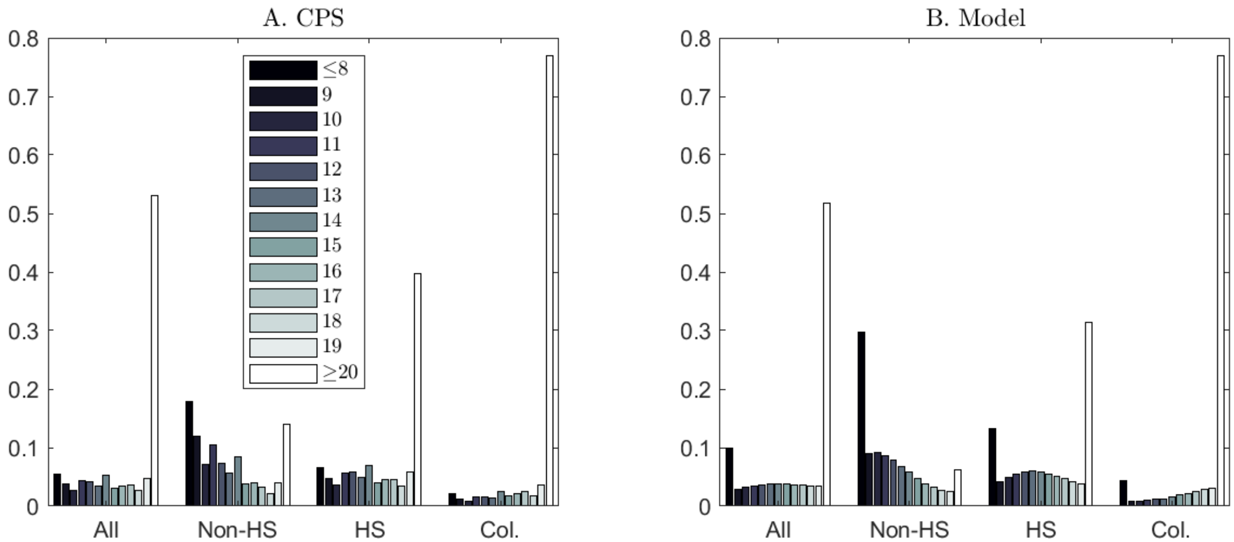


Figure A1: Distribution of wages in model vs. data (CPS)

B Validation of efficiency channels versus recent empirical evidence

The main text describes 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. Recent empirical studies speak directly to these channels: direct effects are measured by Jardim, Long, Plotnick, Van Inwegen, Vigdor, and Wething (2022) and Azar, Huet-Vaughn, Marinescu, Taska, and von Wachter (2023, henceforth AHMTV), spillovers are measured by Engbom and Moser (2022, henceforth EM) and reallocation is measured by Dustmann, Lindner, Schönberg, Umkehrer, and vom Berge (2022, henceforth DLSUB). Two of the studies are from non-U.S. economies due to measurement error concerns for spillovers and lack of comparable reallocation estimates in the U.S. However, with this caveat in mind, we show that the model produces comparable qualitative and quantitative responses.

This implied that the small efficiency gains that we compute in Section 5 are not due to undershooting on any of these mechanisms. Rather, existing reduced form evidence *pointing to* the possibility of efficiency gains can be generated by the theoretical mechanisms suggested in the empirical literature, but nonetheless efficiency gains may be small.

B.1 Direct effects in Seattle

We replicate the disemployment effects of a high $\underline{w}$ on low wage jobs documented following the Seattle minimum wage increase studied in Jardim et. al. (2022).

Empirical setting. Jardim et. al. (2022) study the minimum wage increases in Seattle in 2015 and 2016. These are useful benchmarks as (i) they are minimum wage increases from initially high minimum wages, (ii) the authors have access to hours data which most closely maps into our model concept of n_{ijh} from an efficiency perspective since it is the object that enters production. The authors study two minimum wage increases: *"The minimum wage rose from the state's minimum of \$9.47 to as high as \$11 on April 1, 2015, and again to as high as \$13 on January 1, 2016"* (page 266, and Table 1). The authors compare single-establishment firms in Seattle to those in Washington state, and compute the elasticity of employment in jobs that pay less than \$19 per hour, which account for 63 percent of the workforce (page 269, and Table 2). In Tables 6A and 6B the authors present estimates of causal effects in percent changes on wages and hours. Their results vary across specifications. We summarize them as ranges via their text:

1. **Wages** - We associate the first minimum wage increase with wage effects of 1.1 to 2.2 percent, averaging 1.7 percent, the second increase is associated with a larger 3.0 to 3.9 percent, averaging 3.4 percent wage effect. (page 290)
2. **Hours** - Point estimates for the \$11 period range +0.8 and -2.7 percent, averaging -1.0 percent, the subsequent increase to \$13 is associated with larger reductions between 4.6 and 9.9 percent, averaging -7.0 percent. (page 292-3)

Replication. Our economy is calibrated to 2019, so we first deflate all wages to 2015 levels at 1.55 percent inflation using the 2015-2019 CPI. We take an economy with a $\underline{w}$ of \$9.47 to match the pre-2015 baseline. We then consider a \$1.53 minimum wage increase, corresponding to the first raise, and \$2 minimum wage increase corresponding to the second raise. We keep all jobs of all worker types that had a pre-policy wage less than \$20, to match the 63 percent of employment in the study, which applied a very similar cut-off of \$19. We then compute the percent change in total employment—which corresponds to hours in their study—and the

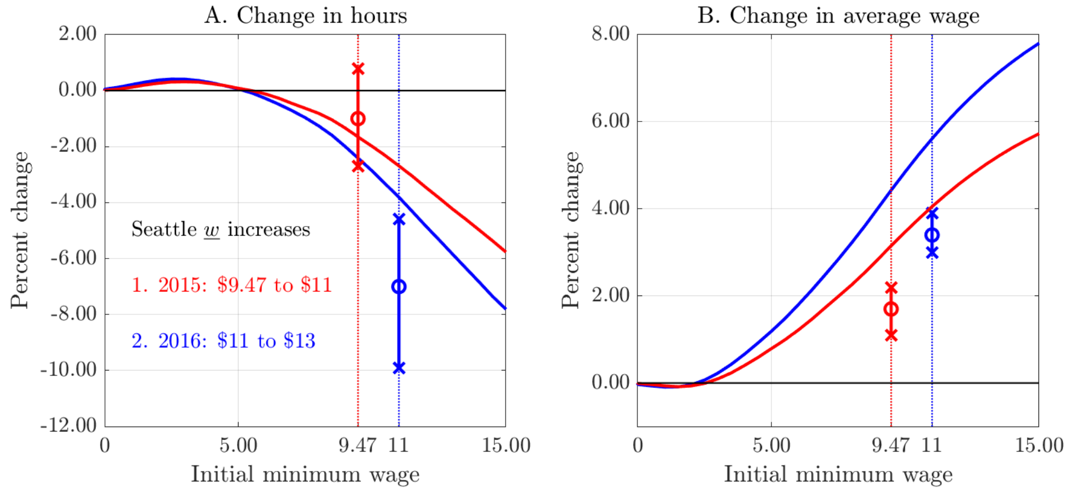


Figure B1: Disemployment effects on low wage employment from high initial minimum wages - Seattle

Notes: **Panel A.** The red line plots percent changes in hours following a $1.53 increase in $\underline{w}$, and the blue line following a $2.00 increase for an initial minimum wage specified on the x-axis. Vertical dashed lined denote the initial minimum wages of $9.47 (red) and $11 (blue). Solid box-whisker lines denote the range of point estimates described by the authors in Jardim et. al. (2022), see text. **Panel B.** Repeats the same exercise as Panel A with average wages.

average wage. A benefit of the model is that we can conduct this for multiple initial minimum wages. We do not recalibrate any other parameters to Seattle data.

Results. Figure B1 presents our results. The vertical lines denote the aforementioned ranges of point estimates and average estimate from the authors. The horizontal axis plots the initial minimum wage. The red line plots percent changes in hours and average wage following a $1.53 increase in $\underline{w}$, and the blue line following a $2.00 increase.

First, consistent with the authors we obtain negative effects on hours and positive effects on wages. Second, the model has similar non-linear employment effects as found in the data. Effects on hours are small following the first increase, and large following the second increase. The model understates the large negative effects on the second increase found in Seattle, but would obtain similar estimates from a $2 increase from $13 to $15 per hour. Third, the increase in average wages is larger for the second increase, in roughly the right proportion to the first increase. However, in levels, our response is only about 1 percentage point larger in the model compared to the authors' empirical estimates.

In summary, these results give us confidence that the non-linearities in the model observed in our welfare exercises are consistent with the data, and kick in at the empirically relevant range of minimum wages, around \$10 to \$13 per hour.

B.2 Direct employment effects in concentrated markets

We also analyze the direct effects of minimum wages and how they vary by market structure. AHMTV highlight the positive effects of minimum wages on employment in high concentration markets (where concentration is measured by the Herfindahl index of employment in a local labor market), and the negative effects of minimum wages on employment in low concentration markets. We further demonstrate that at low levels of the minimum wage, small changes in the minimum wage generate employment increases nationally. These results suggest that minimum wages may reduce markdowns and induce employment expansions, similar to neoclassical frameworks built on Robinson (1933).

Empirical setting. AHMTV compute the response of employment in low wage occupations to changes in state minimum wages, but stratify responses by the concentration of the labor market for each occupation. They estimate statistically significant positive effects in markets in the upper tercile of concentration, and statistically significant negative effects in markets in the lower tercile of concentration. We show that the same results hold in our economy, qualitatively.

Our replication is subject to two caveats. First, AHMTV measure concentration using the Herfindahl of job openings in Burning Glass vacancy data. In a large class of search models with balanced matching, vacancies are proportional to employment. While we do not model job search, we appeal to this intuition and measure concentration using the employment Herfindahl. Second, they restrict their analysis to the retail sector (Stock Clerks, Retail Sales, and Cashiers). The average retail wage is \$16.70[37]. So to align our results with theirs we restrict our analysis to high-school "retail sales" workers in the fourth quintile of earnings whose average wage is \$16.76, and thus maps most closely to AHMTV.

Statistic. Holding aggregates fixed, we increase the minimum wage by ϕ and compute the increase in employment in each market j. Exactly as in AHMTV, we regress the change in market employment $\Delta \log n_j$ on the change in the minimum wage $\Delta \log \underline{w}$, interacted with dummies for DOJ concentration thresholds based on

[37] https://www.bls.gov/oes/current/oes412031.htm, accessed September 2023.

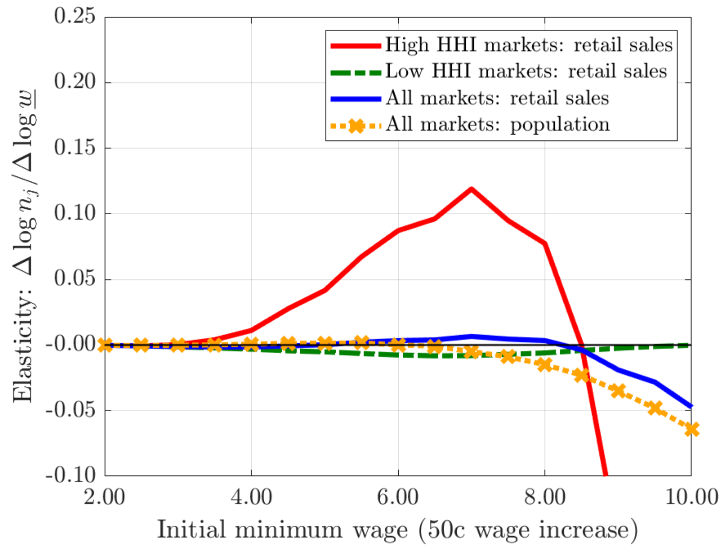


Figure B2: Replication of direct effects by HHI

Notes: Horizontal axis gives the initial minimum wage $\underline{w}_0$. The minimum wage is then increased by 50 cents. Red solid line plots estimated elasticity in high concentration markets ($\widehat{\psi}_L + \widehat{\psi}_H$). Green dashed line plots estimated elasticity in low concentration markets ($\widehat{\psi}_L$). Blue and Orange-Cross lines represent pooled effects for "retail" and the total population, respectively.

the employment Herfindahl (HHI^n):[38]

$$\Delta \log n_j = \psi_L \Delta \log \underline{w} + \psi_H D\Big(HHI_j^n \in [0.25, \infty)\Big) \times \Delta \log \underline{w} + \varepsilon_j.$$

In their sample, the average pre- and post-policy minimum wages are \$7.43 and \$7.83, which we round to $\phi = 50$cents.[39] We use the model to understand heterogeneity by the level of the initial minimum wage, repeating this exercise for initial minimum wages $\underline{w}_0$ between \$2 and \$10 per hour.

Results. Figure B2 plots the estimated coefficients for low ($\widehat{\psi}_L$) and high ($\widehat{\psi}_L + \widehat{\psi}_H$) concentration markets, holding the increase in the minimum wage constant (50c), but varying the initial minimum wage $\underline{w}_0$. For $\underline{w}_0$ consistent with the setting the paper studies—i.e. less than \$8.00 per hour—the model is consistent with AHMTV's key empirical findings. High concentration markets experience positive employment effects (solid red), and low concentration markets experience small negative employment effects (dashed green). Our peak employment elasticity in high concentration markets is 0.12. Our point estimate is about 41% of theirs (=0.12/0.29) and very close to the lower bound of the 95% CI of 0.17 (Ta-

[38] Following the methodology of AHMTV, we average the pre- and post- minimum wage Herfindahl.

[39] We thank the authors for sharing these two statistics with us.

ble 2, Col 2 of AHMTV). Firms in more concentrated markets have more market power, wider markdowns, and hence have larger positive employment gains available in Region II before shrinking in Region III. The expansion of employment in concentrated markets is evidence of direct effects of minimum wages reducing markdowns and inducing employers to expand. In less concentrated markets, firms have initially narrow markdowns and move quickly into Region III, incurring employment losses. Crucially, the positive effects of minimum wages occur in concentrated markets.

Lastly, we run two unconditional regressions to measure the aggregate elasticity of employment:

$$\Delta \log n_j = \psi_{pooled} \Delta \log \underline{w} + \varepsilon_j$$

We estimate ψ_{pooled} for (1) "Clerks" (blue solid), and (2) the overall population (orange crosses). We find broadly similar results: employment expands initially following increases from initially low levels of the minimum wage and then contracts once the initial minimum wage is beyond \$8.00 per hour. The employment expansion at low levels of the minimum wage is, through the lens of our model, due to a reduction in markdowns as firms enter Region II. Overall, low concentration markets dominate the response as they employ the most workers.

Among the positive employment elasticities reported among U.S. studies (see Neumark and Shirley (2022) and Clemens and Strain (2021)), our model's small positive employment elasticities fall within the range reported by the literature.

B.3 Spillovers from competitors' minimum wages

Our next validation exercise is to replicate the spillover effects observed in EM. There are estimates of U.S. spillovers based on survey data, however Autor, Manning, and Smith (2016) argue that measurement error in survey data poses significant issues for inference.[40] EM avoid these issues via administrative data on hours and wages from Brazil. Additionally, we focus on EM since they provide the necessary summary statistics for replication.

Empirical setting. EM compute that in 1996 the minimum wage was 30.3 percent of the median wage. It then increased by 128 percent between 1996 and 2012 (EM,

[40] Autor, Manning, and Smith (2016)'s concern is that measurement error in U.S. survey data can explain the majority of measured spillover effects. See Section IV of Autor, Manning, and Smith (2016).

page 3813). To replicate this experience we solve our economy under a minimum wage of $5.50, which is 30.0 percent of the median wage, then increase it to $12.50 which is a 128 percent increase. We denote these period zero and period one.

Statistic. Let $\overline{p}$ be a reference percentile of the wage distribution, and let $w_{p,t}$ be the percentile p wage in period t. We compute spillovers at p by

$$Spillover_p = \frac{\log(w_{p,1}/w_{\overline{p},1}) - \log(w_{p,0}/w_{\overline{p},0})}{\log(\underline{w}_1/w_{\overline{p},1}) - \log(\underline{w}_0/w_{\overline{p},0})} \tag{B1}$$

By construction $Spillover_{\overline{p}} = 0$. If wages below $\overline{p}$ compress upward, then $Spillover_p > 0$. If wages above $\overline{p}$ compress upward, then $Spillover_p < 0$. EM use a regression framework to obtain estimates of $Spillover_p$, whereas we simply compute $Spillover_p$ non-parametrically via (B1). As shown by (EM, Figure A2), even within the 70th percentile of the earnings distribution more than 80 percent of workers have not completed high school in Brazil.[41] We therefore compute results for non-High school workers.

Results. Figure B3 plots $Spillover_p$ for $p \in [10, 12, \ldots, 90]$ and compares estimates to those from (EM, Figure 4) where the reference percentile is $\overline{p} = 50$.[42] We find very similar qualitative and quantitative patterns of spillovers, with compression far up into the wage distribution. At the 30th percentile, wages compress by 22% in the data versus 35% in the model. By construction the spillover is zero at $\overline{p} = 50$. At the 80th percentile, wages compress by 17% in the data versus 20% in the model. While the US labor market is subject to very different institutions than the Brazilian labor market, we view Figure B3 as a validation of our model's mechanisms on the best available data.

Additional replication. 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.

[41] Another statistic that reflect this is as follows: at the 90th percentile of the earnings distribution, only 10 percent of workers have a college degree.

[42] We compare our results to their IV specification that controls for state-level trends and state fixed effects. This delivers similar results to their specification with state-level fixed effects only.

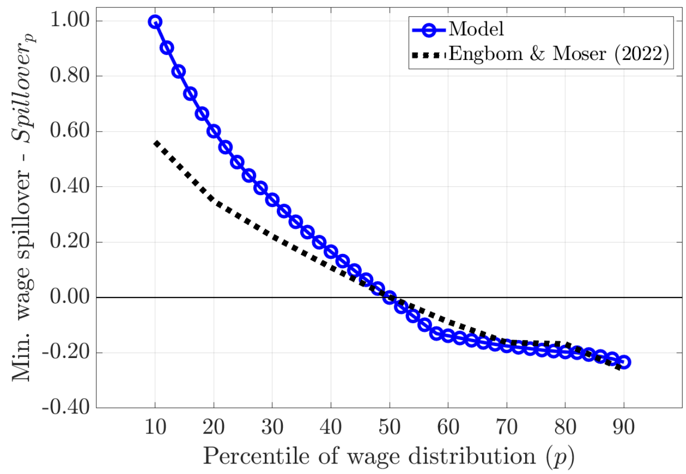


Figure B3: Replication of wage spillovers

Notes: Consistent with the minimum wage in Brazil in 1996, the initial minimum wage is 30 percent of the median wage. Consistent with the minimum wage increase in Brazil from 1996 to 2012, the minimum wage increases by 128 percent. These statistics are reported in Engbom and Moser (2022, page 12). We compare model results to those of Engbom and Moser (2021), Figure 4, under the 'State Fixed Effects plus Trends' OLS specification.

B.4 Reallocation effects of minimum wages

Our final validation exercise replicates DLSUB, *"Reallocation Effects of the Minimum Wage,"* who study the effect of the introduction of a minimum wage in Germany and its impact on the cross-section of workers and firms. In January 2015, a national minimum wage of 8.50 euros was introduced into an environment with no pre-existing minimum wage. This corresponds to \$10.40/hr in 2019 US dollars. The minimum wage introduced in Germany was large: pre-reform, 15 percent of workers earned below 8.50, which was 48 percent of the median wage. The key finding is employment reallocation: small firms exit, and larger more productive firms expand, increasing average firm size.

Empirical setting. DLSUB consider a number of empirical approaches. The one we focus on computes the elasticity of firm characteristics with respect to minimum wage exposure. The authors compute a *Gap* measure: the percent increase in total earnings required to satisfy the new minimum wage, holding employment and hours fixed at their pre-reform level. Let workers be indexed by $\ell \in \{1, \ldots, n\}$. DLSUB define *Gap* using workers' pre-reform hours h_ℓ and wages w_ℓ:

$$Gap := \left[\sum_\ell \max\left\{\underline{w} - w_\ell, 0\right\} h_\ell\right] \Big/ \left[\sum_\ell w_\ell h_\ell\right]$$

The authors group firms by geographic regions r, and regress changes in region outcomes on Gap_r. We focus on two dependent variables studied: (i) total number of operating firms, and (ii) average firm size. Their results are in Table 7, page 54.

Replication. To an economy with no minimum wage, we introduce a minimum wage of \$9.85/hr. This is relatively low, but equals 48 percent of the pre-reform median wage. The empirical setting is a national reform, so we solve the pre- and post-reform economy in general equilibrium. The regions considered in DLSUB, comprise all industries in multiple commuting zones and rural areas. These are much larger than markets j in our model. We therefore treat our whole economy as one region, which generates a single Gap measure directly comparable to theirs:

$$Gap = \left[\sum_h \int \sum_i \max\{\underline{w} - w_{ijh}, 0\} n_{ijh}\, dj\right] \Big/ \left[\sum_h \int \sum_i w_{ijh} n_{ijh}\, dj\right]. \tag{B2}$$

To compute the elasticity of variable x with respect to Gap, we divide economy-wide $\Delta \log x$ by Gap.

Results. Figure B4 gives the results.[43] We plot results for a range of minimum wages $\underline{w}_1$, indexed by the ratio of $\underline{w}_1$ to the pre-reform median wage w_0^{p50}. The vertical line marks the $\underline{w}_1 / w_0^{p50} = 0.48$ corresponding to DLSUB.

Consistent with the new reallocation facts in DLSUB, Panel A shows that reallocation causes average firm size to grow and Panel B shows that small firms exit. In the model all firms still operate due to decreasing returns and since n_{ijh} is continuous it can go below one (recall Figure 1D). To compare our model to DLSUB, we classify a firm as 'operating' when their employment is above one worker.

The model's elasticity of average firm size with respect to minimum wage exposure is positive and in line with the data (Figure B4A). The increase in average firm size represents reallocation, and is moderated at larger minimum wage increases due to firms shrinking in Region III, consistent with positive gains from reallocation being limited to small minimum wage increases. The elasticity of the number of operating firms with respect to Gap is negative and thus correctly signed, but more responsive compared to the data (Figure B4AB.

Interpretation. One of the key take-aways of DLSUB is that minimum wage increases have heterogeneous effects across firms. Low productivity firms exit, but

[43] There are two sets of the authors' results: 'Data 1' and 'Data 2'. Both feature controls that account for observable regional differences (e.g. average age) and region specific trends in the moments. 'Data 2' additionally interacts these trends with year fixed effects.

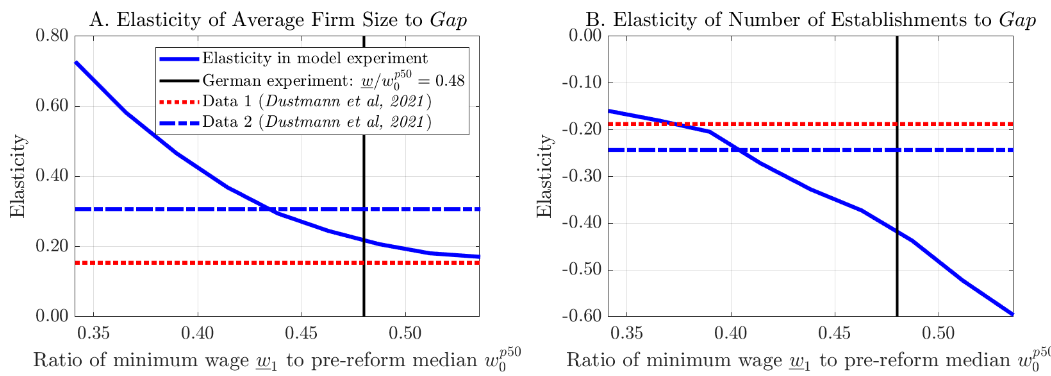


Figure B4: Replication of DLSUB (2021) - Reallocation effects of minimum wages

Notes: Corresponding data estimates for "Data 1" and "Data 2" are respectively taken from p.54 of DLSUB, Table 7, Columns (2) [regional controls and region specific linear trend] and (4) [regional controls interacted with year fixed effects]. The solid blue line plots the elasticity of the relevant moment to the minimum wage *Gap*, computed as in equation (B2). The horizontal axis plots the minimum wage in the policy experiment simulated in the model as a fraction of the pre-reform median wage in the model.

their workers do not move out of the labor market. Jobs which existed due to the small amount of market power at these low productivity firms are destroyed, but workers are reallocated to larger, more productive firms. This can improve allocative efficiency. Our model generates dynamics consistent with these observations.

C Robustness exercises

C.1 Varying Frisch elasticity φ

Our main results are robust to the Frisch elasticity of labor supply. We consider two values of φ either side of the baseline value of 0.62. These values are informed by our exercise in Appendix E using data from Golosov et. al. (2021). Their results imply larger φ for high income households (lower MPC, higher MPE) than low income households (higher MPC, lower MPE). We consider values that match data for both groups: $\varphi \in \{0.30, 0.86\}$.[44] We recalibrate all other 'shifter' parameters to match data in Table A1. Appendix Tables C1 and C2 show that levels of φ have essentially zero effect on our calculations.

[44] This range subsumes the range used by the Congressional Budget Office when modeling policy, which is around 0.30 to 0.53. See the following (link).

A. Maximize utilitarian welfare, $\Lambda_\pi(\underline{w})$		**Alt. Economy**	**Baseline**
Optimal minimum wage ($)	$\underline{w}^*$	10.93	10.95
Welfare (%)	$\Lambda_\pi(\underline{w}^*)$	2.77	2.79
Aggregate efficiency (%)	$AE(\underline{w}^*)$	-0.13	-0.07
B. Maximize aggregate efficiency, $AE(\underline{w})$		**Alt. Economy**	**Baseline**
Optimal minimum wage ($)	$\underline{w}^{*,AE}$	7.01	7.35
Welfare (%)	$\Lambda_\pi(\underline{w}^{*,AE})$	1.40	1.55
Aggregate efficiency (%)	$AE(\underline{w}^{*,AE})$	0.08	0.09

C. Moments in alternative economy		**NHS**	**HS**	**C**	**O**
Relative population* (%)	$\pi_h / \sum \pi_h$	13.22	53.70	26.08	7.00
Average earnings per hour*, (C=1)		0.40	0.59	1.00	
Share of aggregate labor income* (CPS and SCF, %)		3.0	38.5	46.2	12.4
Binding at $15, all* (%)		—— 30.57 ——			

Table C1: Robustness exercise - $\varphi = 0.30$

A. Maximize utilitarian welfare, $\Lambda_\pi(\underline{w})$		**Alt. Economy**	**Baseline**
Optimal minimum wage ($)	$\underline{w}^*$	10.97	10.95
Welfare (%)	$\Lambda_\pi(\underline{w}^*)$	2.80	2.79
Aggregate efficiency (%)	$AE(\underline{w}^*)$	-0.03	-0.07
B. Maximize aggregate efficiency, $AE(\underline{w})$		**Alt. Economy**	**Baseline**
Optimal minimum wage ($)	$\underline{w}^{*,AE}$	7.54	7.35
Welfare (%)	$\Lambda_\pi(\underline{w}^{*,AE})$	1.64	1.55
Aggregate efficiency (%)	$AE(\underline{w}^{*,AE})$	0.10	0.09

C. Moments in alternative economy		**NHS**	**HS**	**C**	**O**
Relative population* (%)	$\pi_h / \sum \pi_h$	13.22	53.70	26.08	7.00
Average earnings per hour*, (C=1)		0.40	0.59	1.00	
Share of aggregate labor income* (CPS and SCF, %)		3.0	38.5	46.2	12.4
Binding at $15, all* (%)		—— 30.57 ——			

Table C2: Robustness exercise - $\varphi = 0.86$

C.2 Heterogeneous region calibration

We split our economy into three separate regions, denoted r and consider a separate household type for each region.[45] We calibrate each region to data from three sets of US states, grouped by median household income. Each region contains approximately one third of the civilian labor force.[46] Across regions, we keep

[45]We make the simplifying assumption that labor is immobile across regions. Capital and consumption goods are traded at the same rental rate and price across all regions.

[46]States are allocated to regions as followed, ordered by 2019 median household income within each region. Low income states: MS, LA, NM, WV, AR, KY, AL, TN, GA, FL, OK, MT, MS, NC, SC, MI, SD. Medium income states: OH, WY, ID, IA, ME, IN, WI, TX, ND, RI, PA, AZ, NV, NY, CO, NE, KS, DE, VT. High income states: IL, OR, CA, AK, VA, MN, WA, UT, NH, CT, MA, NJ, HI, DC, MD.

A. Maximize utilitarian welfare, $\Lambda_\pi(\underline{w})$		Alt. Economy		Baseline	
Optimal minimum wage ($)	$\underline{w}^*$	10.68		10.95	
Welfare (%)	$\Lambda_\pi(\underline{w}^*)$	2.91		2.79	
Aggregate efficiency (%)	$AE(\underline{w}^*)$	-0.13		-0.07	
B. Maximize aggregate efficiency, $AE(\underline{w})$		**Alt. Economy**		**Baseline**	
Optimal minimum wage ($)	$\underline{w}^{*,AE}$	6.76		7.35	
Welfare (%)	$\Lambda_\pi(\underline{w}^{*,AE})$	1.42		1.55	
Aggregate efficiency (%)	$AE(\underline{w}^{*,AE})$	0.10		0.09	
C. Moments in alternative economy		**NHS**	**HS**	**C**	**O**
Relative population* (%)	$\pi_h / \sum \pi_h$	13.05	57.19	22.75	7.00
Average earnings per hour*, (C=1)		0.42	0.62	1.00	
Share of aggregate labor income* (CPS and SCF, %)		3.2	44.6	39.9	12.3
Binding at $15, all* (%)		—— 34.21 ——			

Table C3: Robustness exercise - States in lowest tercile of income

some preference and technology parameters the same, as well as the distribution of number of firms in a market: $\{\beta, \theta, \eta, \delta, \alpha, \gamma, G(M_j)\}$. Region r shifters $(\overline{\varphi}_r, \overline{Z}_r)$, $\{\overline{\varphi}_{kr}, \xi_{kr}\}_{k=1,r=1}^{K,R}$ and measures $\{\pi_{kr}\}_{k=1,r=1}^{K,R}$, are chosen to match CPS data from each region, following Table A1. Since the SCF does not identify an individual's state, we impose three further restrictions across regions. We keep constant (i) the ratio of household capital to labor income,[47] (ii) the fraction of households that are owners,[48] (iii) average firm size, which determines $\overline{\varphi}_r$.

Tables C3 and C4 show that relative to High income states, Low income states have significantly lower wages (last row). A $15 minimum wage binds for 34% of low income state workers, and only 27% for high income state workers.

The greater rationing among low income states puts downward pressure on both the Utilitarian and efficiency maximizing minimum wages. The high income states have a marginally higher Utilitarian minimum wage (+$0.91) and a marginally higher efficiency maximizing minimum wage (+$1.17). However, aggregate efficiency gains are still less than 0.10% in both.

We repeat this exercise for Mississippi (MS) in Table C5, recalibrating to match 41.3% of workers below $15 an hour in MS. The optimal minimum wage falls by

[47]Since other parameters change, we recalibrate the share parameters $\{\kappa_{kr}\}_{k=1,r=1}^{K,R}$ to match the benchmark targets.

[48]For example, if Region A has 37% of workers with a college degree, and Region B has 29%, then in both Region A and Region B we maintain that 7% of households are *college-owners* (Table A1) and set the share of households that are *college-workers* to 30% in Region A and 22% in Region B.

A. Maximize utilitarian welfare, $\Lambda_\pi(\underline{w})$		Alt. Economy		Baseline	
Optimal minimum wage ($)	$\underline{w}^*$	11.59		10.95	
Welfare (%)	$\Lambda_\pi(\underline{w}^*)$	2.80		2.79	
Aggregate efficiency (%)	$AE(\underline{w}^*)$	-0.06		-0.07	
B. Maximize aggregate efficiency, $AE(\underline{w})$		**Alt. Economy**		**Baseline**	
Optimal minimum wage ($)	$\underline{w}^{*,AE}$	7.93		7.35	
Welfare (%)	$\Lambda_\pi(\underline{w}^{*,AE})$	1.62		1.55	
Aggregate efficiency (%)	$AE(\underline{w}^{*,AE})$	0.09		0.09	
C. Moments in alternative economy		**NHS**	**HS**	**C**	**O**
Relative population* (%)	$\pi_h / \sum \pi_h$	13.33	50.34	29.33	7.00
Average earnings per hour*, (C=1)		0.40	0.57	1.00	
Share of aggregate labor income* (CPS and SCF, %)		3.1	33.6	51.0	12.2
Binding at $15, all* (%)			—— 27.26 ——		

Table C4: Robustness exercise - States in highest tercile of income

A. Maximize utilitarian welfare, $\Lambda_\pi(\underline{w})$		Alt. Economy		Baseline	
Optimal minimum wage ($)	$\underline{w}^*$	9.24		10.95	
Welfare (%)	$\Lambda_\pi(\underline{w}^*)$	2.75		2.79	
Aggregate efficiency (%)	$AE(\underline{w}^*)$	-0.09		-0.07	
B. Maximize aggregate efficiency, $AE(\underline{w})$		**Alt. Economy**		**Baseline**	
Optimal minimum wage ($)	$\underline{w}^{*,AE}$	5.96		7.35	
Welfare (%)	$\Lambda_\pi(\underline{w}^{*,AE})$	1.42		1.55	
Aggregate efficiency (%)	$AE(\underline{w}^{*,AE})$	0.10		0.09	
C. Moments in alternative economy		**NHS**	**HS**	**C**	**O**
Relative population* (%)	$\pi_h / \sum \pi_h$	15.98	63.01	14.01	7.00
Average earnings per hour*, (C=1)		0.45	0.68	1.00	
Share of aggregate labor income* (CPS and SCF, %)		5.4	52.9	27.8	13.9
Binding at $15, all* (%)			—— 41.43 ——		

Table C5: Robustness exercise - Mississippi

$1.71 due to the stronger degree of rationing and greater share of non-highschool workers. The efficiency maximizing minimum wage also falls by $1.40 for similar reasons. As in our other regional exercises, firm heterogeneity and rationing of the lowest wage workers mutes the efficiency gains from minimum wages.

Overall we note that efficiency maximizing minimum wages are below $8 across these exercises, and efficiency gains are less than 0.10%.

C.3 Fixed capital: Short run vs. long run

In comparing steady-states we are implicitly studying the long-run effects of the minimum wage. Our theory suggests a smaller optimal minimum wage in the short-run if the cost of labor increases but the level and distribution of capital across workers in each firm is slow to adjust. If we assume maximal stickiness in reallocation of capital across-workers within-firm (fixed capital) a minimum wage causes exit, but we find these effects are quantitatively small. When capital is fixed, the optimal minimum wage under Utilitarian weights declines by 80 cents (Table C2). With sharper decreasing returns in the short-run, the range of productivity for which firms are in Region II shrinks (Figure C1), reducing potential efficiency gains. Quantitatively, short- and long-run elasticities in our model are similar which is reassuring for our mapping to empirical studies of short-run changes.

We provide the theory and details for the short vs. long-run exercise above. We increase the minimum wage but keep firm-worker specific installations of capital fixed at the allocation $\overline{k}_{ijh}$ under a zero minimum wage. Firm profits from each type are as follows:

$$\pi_{ijh} = \overline{Z}\xi_h \left(z_{ij}\overline{k}_{ijh}^{(1-\gamma)\alpha} \right) n_{ijh}^{\gamma\alpha} - w_{ijh}n_{ijh} - R\overline{k}_{ijh}.$$

First, with fixed capital, the production function has sharper decreasing returns in labor: $\gamma\alpha < \widetilde{\alpha}$. Second, firms face overheard costs of pre-installed capital, $R\overline{k}_{ijh}$, which will cause termination of non-profitable jobs at high minimum wages. We therefore add an endogenous margin of operation into the solution of the model.[49] Third, equilibrium conditions are as before, minus the capital demand condition. Capital supply remains infinitely elastic at $R = 1/\beta + (1-\delta)$, but demand is pinned down at $\overline{K}(\underline{w}) = \sum_h \int \sum_i \chi_{ijh}(\underline{w})\overline{k}_{ijh}\, dj$, where $\chi_{ijh}(\underline{w}) \in \{0,1\}$ indicates whether the firm operates worker-type-k capital in equilibrium under minimum wage $\underline{w}$.

Figure C1 characterizes the mechanism behind a lower optimal minimum wage in this environment. Panel A considers a firm in an economy without a minimum

[49] Market-by-market we first assume that all firms enter, and then solve the Nash equilibrium of the market and general equilibrium of the economy. We then compute firm-type profits π_{ijh}, which account for fixed capital costs. If any firm has profits $\pi_{ijh} < 0$, we drop the lowest productivity firm in the market and then solve the market equilibrium again. With fewer firms, labor market power of the remaining firms increases, which increases profits, hence the need to remove only one firm at a time. We continue in this way until we reach a Cournot Nash equilibrium: no firm with shut-down jobs wishes to re-open them given competitor's operation and intensive margin labor decisions.

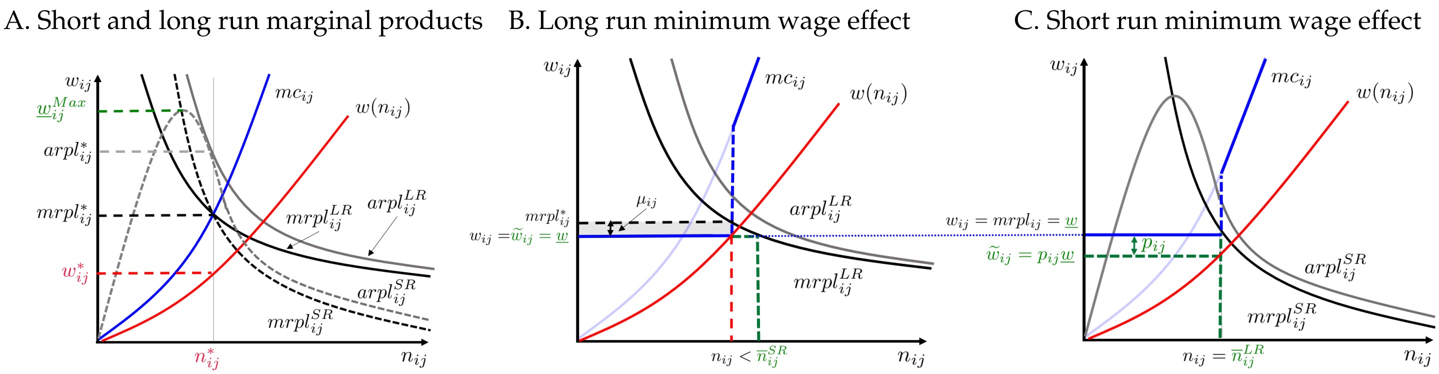


Figure C1: Partial equilibrium theory of minimum wage in the short-run

wage, where capital is fixed at the allocation consistent with long-run employment n^*_{ij}. Short-run marginal and average products coincide with long-run values at this point. Away from n^*_{ij}, short-run $mrpl^{SR}_{ij}$ is steeper due to sharper decreasing returns with fixed capital: if $n_{ij} > n^*_{ij}$, then $mrpl^{SR}_{ij} < mrpl^{LR}_{ij}$. With fixed overhead capital, the $arpl^{LR}_{ij}$ goes to zero as n_{ij} goes to zero and overhead per worker explodes. The peak in $arpl^{SR}_{ij}$ intersects $mrpl^{SR}_{ij}$ and gives the maximum $\underline{w}$ the firm could afford and still operate type-k capital: $\underline{w}^{Max}_{ij}$. At $\underline{w} > \underline{w}^{Max}_{ij}$, equating $\underline{w} = mrpl^{SR}_{ij}$ would imply $arpl^{SR}_{ij} < \underline{w}$ and shutdown is optimal.

Panels B and C show how these differences constrain the positive efficiency gains from narrowing $\widetilde{\mu}_h$. Take the firm in Panel A, in the long run, at the minimum wage pictured in Panel B, the firm is in Region II: employment is non-rationed ($n_{ij} < \overline{n}^{SR}_{ij}$), and wages are a narrower markdown on $mrpl^{LR}_{ij}$. A small increase in the minimum wage *increases employment* and *narrows shadow markdowns*. Panel C considers the short run, at the same minimum wage. The lower $mrpl^{SR}_{ij}$ places the firm in Region III, where employment is rationed. A small increase in the minimum wage *decreases employment* and *widens shadow markdowns*. In the short run, the range of $\underline{w}$ over which firms are in Region II is smaller. This constrains the efficiency gains from improvements in $\widetilde{\mu}_h$.

Table C2 reports the results from fixing capital. The short-run optimal minimum wage under Utilitarian weights declines by about 80 cents. Likewise, the aggregate efficiency maximizing minimum wage is roughly 20 cents lower. Thus with sharper decreasing returns in the short-run, there is a smaller range of productivity for which firms are in Region II and efficiency gains decline.

A. Maximize utilitarian welfare, $\Lambda_\pi(\underline{w})$		**Alt. Economy**	**Baseline**
Optimal minimum wage ($)	$\underline{w}^*$	10.12	10.95
Welfare (%)	$\Lambda_\pi(\underline{w}^*)$	2.33	2.79
Aggregate efficiency (%)	$AE(\underline{w}^*)$	-0.03	-0.07
B. Maximize aggregate efficiency, $AE(\underline{w})$		**Alt. Economy**	**Baseline**
Optimal minimum wage ($)	$\underline{w}^{*,AE}$	7.08	7.35
Welfare (%)	$\Lambda_\pi(\underline{w}^{*,AE})$	1.42	1.55
Aggregate efficiency (%)	$AE(\underline{w}^{*,AE})$	0.08	0.09

C. Moments in alternative economy		**NHS**	**HS**	**C**	**O**
Relative population* (%)	$\pi_h / \sum \pi_h$	13.22	53.70	26.08	7.00
Average earnings per hour*, (C=1)		0.40	0.59	1.00	
Share of aggregate labor income* (CPS and SCF, %)		3.0	38.5	46.2	12.4
Binding at $15, all* (%)		—— 30.57 ——			

Figure C2: Robustness - Fixed capital / Short-run

C.4 Labor-labor substitution

Despite the additive nature of our production function, type-level decreasing returns implies different types of labor are not perfect substitutes. As shown in equation (9), the elasticity of substitution between different education groups is $(1-\alpha(1-\gamma))/(1-\alpha)$. Baseline $\alpha = 0.94$ and $\gamma = 0.81$ implies an elasticity of 13.7, which is high relative to the literature.[50] We consider an alternative calibration with $\alpha = 0.70$ which delivers an elasticity of substitution of 2.9, close to the value estimated by Acemoglu and Autor (2011) which extended Katz and Murphy (1992) through 2008. Qualitatively, as in the above short-run exercise, a lower α steepens the labor demand curve, reducing the range over which a firm will be found in Region II, choking off the *Direct effect*. The efficiency maximizing minimum wage falls to $6.82 and the welfare gains from improvements in efficiency fall slightly from 0.09% to 0.08%.

C.5 Capital-labor substitution

Our benchmark model features an elasticity of substitution between capital and labor of 1.0. Prominent existing studies estimate lower elasticities around 0.7 (Oberfield and Raval, 2021) and high elasticities around 1.2 (Karabarbounis and Neiman, 2014), with the majority of studies pointing to estimates less than 1.0 (Gechert,

[50] See for example, Katz and Murphy (1992), Card and Lemieux (2001), and Acemoglu and Autor (2011) which falls in the range of 1.5 to 2.9.

A. Maximize utilitarian welfare, $\Lambda_\pi(\underline{w})$		Alt. Economy		Baseline	
Optimal minimum wage ($)	$\underline{w}^*$	10.99		10.95	
Welfare (%)	$\Lambda_\pi(\underline{w}^*)$	2.56		2.79	
Aggregate efficiency (%)	$AE(\underline{w}^*)$	-0.01		-0.07	
B. Maximize aggregate efficiency, $AE(\underline{w})$		**Alt. Economy**		**Baseline**	
Optimal minimum wage ($)	$\underline{w}^{*,AE}$	6.82		7.35	
Welfare (%)	$\Lambda_\pi(\underline{w}^{*,AE})$	1.47		1.55	
Aggregate efficiency (%)	$AE(\underline{w}^{*,AE})$	0.08		0.09	
C. Moments in alternative economy		**NHS**	**HS**	**C**	**O**
Relative population* (%)	$\pi_h / \sum \pi_h$	13.22	53.70	26.08	7.00
Average earnings per hour*, (C=1)		0.40	0.59	1.00	
Share of aggregate labor income* (CPS and SCF, %)		3.0	38.5	46.2	12.4
Binding at $15, all* (%)		—— 30.61 ——			

Figure C3: Robustness - Labor-labor substitution

Havranek, Irsova, and Kolcunova, 2022). Our baseline value of a unitary elasticity is within this range. It is important to note that lower elasticities do little to our main result, namely that the efficiency maximizing minimum wage is relatively small. The above short-run exercise incorporates an extreme elasticity of substitution of zero and finds the efficiency maximizing minimum wage is unaltered.

D Proofs

We characterize the equilibrium of a simple oligopsony economy without capital or worker heterogeneity in order to ease exposition. Firms and market structure are identical to the economy in the text, except they do not rent capital. Workers have linear preferences over consumption to simplify the labor supply system. The proofs generalize to economies that do not make these assumptions.

Preferences. Preference over consumption are linear with an additive disutility of supplying labor identical to the main text.

$$\mathcal{U} = C - \frac{N^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} \quad , \quad N := \left[\int_0^1 n_j^{\frac{\theta+1}{\theta}}\, dj\right]^{\frac{\theta}{\theta+1}} \quad , \quad n_j := \left[\sum_{i=1}^{M_j} n_{ij}^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}} \quad , \quad \eta \geq \theta. \tag{D1}$$

Labor market competition. As in the text, firms take actions taking their competitors' employment as given. That is they Cournot compete, and understand that they influence market-level outcomes (i.e. firms are oligopsonists). Actions consist of choosing their own quantity of employment, wage and rationing constraint. Labor market j is infinitesimal with respect to other labor markets in the economy, so firms take quantities and wages outside of their labor market as given.

Household problem. The household takes rationing constraints $\{\overline{n}_{ij}\}$, wages $\{w_{ij}\}$ and profits Π as given (which are accordingly omitted from the optimization problem below). The household chooses employment $\{n_{ij}\}$ at each firm ij to maximize:

$$\max_{\{n_{ij}\}_{i\in\{0,M_j\},j\in[0,1]}} \int_0^1 \sum_{i=1}^{M_j} w_{ij}n_{ij}dj - \frac{N^{1+1/\varphi}}{1+1/\varphi} \tag{D2}$$

subject to $n_{ij} \leq \overline{n}_{ij}$ for all $i \in \{1,\ldots,M_j\}$ and $j \in [0,1]$. Let ν_{ij} be the multiplier on the rationing constraint. The following optimality conditions characterizes the labor supply decision of the household:

$$w_{ij} = \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} + \nu_i \quad , \quad \nu_i(\overline{n}_i - n_i) = 0. \tag{D3}$$

We can combine the conditions in equation (D3) to obtain *the inverse labor supply schedule*, which equates the wage to the marginal disutility of labor:

$$w\Big(n_{ij},\overline{n}_{ij},n_j,N\Big) = \begin{cases} \left(\frac{n_{ijt}}{n_{jt}}\right)^{\frac{1}{\eta}} \left(\frac{n_{jt}}{N_t}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} & , n_{ijt} \in [0,\overline{n}_{ijt}) \\ \in \left[\left(\frac{\overline{n}_{ijt}}{n_{jt}}\right)^{\frac{1}{\eta}} \left(\frac{n_{jt}}{N_t}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}, \infty\right) & , n_{ijt} = \overline{n}_{ijt} \end{cases} \tag{D4}$$

Note that this does not directly depend on the minimum wage. This is a correspondence at $\overline{n}_{ij}$. Given any wage greater than the marginal disutility of labor at $\overline{n}_{ij}$, *in red*, the household will supply $\overline{n}_{ij}$. Note, also, that the marginal disutility of labor at any firm doesn't depend on the *rationing constraints* at other firms, but simply the labor employed at other firms. This is for the standard reason that the first order condition for labor at firm i is a *partial* derivative.

Firm problem. Firm i in market j takes as given local competitors' employment levels $\{n_{-ij}\}$ the aggregate employment index N and chooses its (i) wage w_{ij}, (ii) employment n_{ij}, and (iii) rationing constraint $\overline{n}_{ij}$ in order to maximize profits. The firm is constrained by (a) the minimum wage $w_{ij} \geq \underline{w}$, (b) its rationing constraint $n_{ij} \leq \overline{n}_{ij}$, and (c) the inverse labor supply schedule (D5). Therefore the firm problem is given by,

$$\max_{\overline{n}_{ij},n_{ij},w_{ij}} z_{ij}n_{ij}^{\alpha} - w_{ij}n_{ij} \quad \text{subject to} \quad w_{ij} \geq \underline{w} \ , \ n_{ij} \leq \overline{n}_{ij} \ , \ w_{ij} = w\left(n_{ij},\overline{n}_{ij},n_j,N\right). \tag{D5}$$

The firm understands, directly, $\partial w\left(n_{ij},\overline{n}_{ij},n_j,N\right)/\partial n_{ij} \neq 0$ and, indirectly, via $\partial n_j/\partial n_{ij} \neq 0$ (equation D1), yielding *oligopsonistic* behavior.

Equilibrium. Given a minimum wage $\underline{w}$, an *oligopsonistic Nash-Cournot* equilibrium is (i) a household inverse labor supply curve $w\left(n_{ij},\overline{n}_{ij},n_j,N\right)$, (ii) wages $\left\{w_{ij}\right\}$, (iii) quantities of labor $\left\{n_{ij}\right\}$, (iv) rationing constraints $\left\{\overline{n}_{ij}\right\}$, (v) profits Π, and (vi) aggregate employment index N and market level employment indices $\left\{n_j\right\}$ such that (1) given wages $\left\{w_{ij}\right\}$, rationing constraints $\left\{\overline{n}_{ij}\right\}$, and profits Π, household optimization implies the inverse labor supply curve $w\left(n_{ij},\overline{n}_{ij},n_j,N\right)$, (2) for every firm *i* in market *j*: given competitor employment $\left\{n_{-ij}\right\}$, the aggregate employment index N, and the household inverse labor supply curve, firm *ij*'s optimization yields rationing constraint $\overline{n}_{ij}$, wage w_{ij} and employment n_{ij}, (3) firm employment decisions are consistent with the aggregate and market employment indices, N, $\left\{n_j\right\}$, as well as profits, Π, and (4) markets clear $w_i = w\left(n_{ij},\overline{n}_{ij},n_j,N\right)\ \forall i$.

The remainder of the appendix provides detailed derivations of (1) the firm's perceived inverse labor supply curve and (2) optimal rationing constraint.

(1) Perceived labor supply curve. We proceed via three Lemmas.

Lemma 0 - *Given competitor employment* $\left\{n_{-ij}\right\}$ *, competitor rationing constraints* $\left\{\overline{n}_{-ij}\right\}$ *are payoff irrelevant for firm ij.*

Proof: $\{\overline{n}_{-ij}\}$ do not enter the Cournot oligopsony firm problem. ■

Lemma 1 - *Consider some level of employment* $n_{ij} \leq \overline{n}_{ij}$*. Given competitor employment* $\left\{n_{-ij}\right\}$ *, a firm would never pay a wage that is greater than the lowest legal wage necessary to deliver* n_{ij}*.*

Proof: Conditional on n_{ij} and $\left\{n_{-ij}\right\}$, profits are strictly decreasing in w_{ij}. ■

Lemma 2 - *Consider some level of employment* $n_{ij} \leq \overline{n}_{ij}$*. Given competitor employment* $\left\{n_{-ij}\right\}$*, the lowest legal wage that delivers* n_{ij}*, is given by*

$$\max\left\{\underline{w}, \min\left\{w_L\left(n_{ij},n_j,N\right), w\left(n_{ij},\overline{n}_{ij},n_j,N\right)\right\}\right\} \qquad (*)$$

where $w_L\left(n_{ij},n_j,N\right) = \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}}\left(\frac{n_j}{N}\right)^{\frac{1}{\theta}}N^{\frac{1}{\varphi}}$.

Proof: Given competitor employment $\left\{n_{-ij}\right\}$, the mapping from wages to employment is one-to-one except when $\overline{n}_i = n_i$. In that case $\min\left\{w_L\left(n_{ij},n_j,N\right), w\left(n_{ij},\overline{n}_{ij},n_j,N\right)\right\}$ is the lowest wage that delivers $n_{ij} = \overline{n}_{ij}$ employees. This wage may not be legal.

The lowest legal wage that delivers $n_{ij} = \overline{n}_{ij}$ employees is therefore given by $(*)$. ■

Lemma 2 maps employment to legal wages. We call this mapping the firm's *perceived inverse labor supply curve*, which is the inverse labor supply curve that the firm faces *conditional on* a choice $\overline{n}_{ij}$, and also taking account of the minimum wage. Given competitor employment $\{n_{-ij}\}$ and substituting the firms choice of w_{ij} conditional on choices of $(n_{ij}, \overline{n}_{ij}, n_j)$, we can write the firm's problem as

$$\max_{\overline{n}_{ij}, n_{ij}} z_{ij} n_{ij}^{\alpha} - w_{ij} n_{ij} \quad , \quad \text{subject to}$$

$$w_{ij} = w^p\left(n_{ij}, \overline{n}_{ij}, n_j, N\right) = \begin{cases} \max\left\{\underline{w}, \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}\right\} & , \quad n_{ij} \in \left[0, \overline{n}_{ij}\right) \\ \max\left\{\underline{w}, \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}\right\} & , \quad n_{ij} = \overline{n}_{ij} \end{cases}$$

Using the monotonicity of $\left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}$ in n_{ij},[51] we know that the *highest wage possible* for $n_{ij} \in \left[0, \overline{n}_{ij}\right]$ is at $\overline{n}_{ij}$. If $\left(\frac{\overline{n}_{ij}}{N}\right)^{\frac{1}{\eta}} N^{\frac{1}{\varphi}}$ is less than $\underline{w}$, then it must be the case that $w\left(n_{ij}, \overline{n}_{ij}, n_j, N\right) = \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} < \underline{w}$ for all $n_{ij} \in \left[0, \overline{n}_{ij}\right]$. Define the function $n_j\left(n_{ij}\right)$ as follows:

$$n_j\left(n_{ij}\right) := \left[n_{ij}^{\frac{\eta}{\eta+1}} + \sum_{k \neq i}^{M_j} n_{kj}^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}}.$$

Using this, and given $\overline{n}_{ij}$, we can write the perceived labor supply curve on $n_{ij} \in \left[0, \overline{n}_{ij}\right]$ as follows

$$w^p\left(n_{ij}, \overline{n}_{ij}, n_j, N\right) = \begin{cases} \underline{w} & \text{if } \underline{w} > \left(\frac{\overline{n}_{ij}}{n_j\left(\overline{n}_{ij}\right)}\right)^{\frac{1}{\eta}} \left(\frac{n_j\left(\overline{n}_{ij}\right)}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} \\ \max\left\{\underline{w}, \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}\right\} & \text{if } \underline{w} \leq \left(\frac{\overline{n}_{ij}}{n_j\left(\overline{n}_{ij}\right)}\right)^{\frac{1}{\eta}} \left(\frac{n_j\left(\overline{n}_{ij}\right)}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}. \end{cases}$$

Note that the perceived labor supply curve is not a function of w_{ij}. Given competitor employment $\{n_{-ij}\}$, the Cournot firm problem becomes,

$$\max_{\overline{n}_{ij}, n_{ij}} z_{ij} n_{ij}^{\alpha} - w_{ij} n_{ij}$$

[51] Note $w\left(n_{ij}, \{n_{-ij}\}\right) = \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} = n_{ij}^{\frac{1}{\eta}} n_j^{\frac{1}{\theta} - \frac{1}{\eta}} N^{\frac{1}{\varphi} - \frac{1}{\theta}}$ and that $\partial n_j / \partial n_{ij} > 0$ and $\eta > \theta$. Therefore, given competitor employment $\{n_{-ij}\}$, $\partial w\left(n_{ij}, \{n_{-ij}\}\right) / \partial n_{ij} > 0$.

subject to $n_{ij} \leq \overline{n}_{ij}$ and the *perceived inverse labor supply curve* defined over $[0, \overline{n}_{ij}]$:

$$w^p\left(n_{ij}, \overline{n}_{ij}, n_j, N\right) = \begin{cases} \underline{w} & \text{if } \underline{w} > \left(\frac{\overline{n}_{ij}}{n_j(\overline{n}_{ij})}\right)^{\frac{1}{\eta}} \left(\frac{n_j(\overline{n}_{ij})}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} \\ \max\left\{\underline{w}, \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}\right\} & \text{if } \underline{w} \leq \left(\frac{\overline{n}_{ij}}{n_j(\overline{n}_{ij})}\right)^{\frac{1}{\eta}} \left(\frac{n_j(\overline{n}_{ij})}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}. \end{cases}$$

(2) Optimal rationing constraint. Consider the case of $\overline{n}_{ij} = \infty$. Given competitor employment $\{n_{-ij}\}$, the firm solves the following problem:

$$\max_{n_{ij}} z_{ij} n_{ij}^{\alpha} - \max\left\{\underline{w}, \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}\right\} n_{ij}.$$

We can partition n_{ij} into two sets. Let $\widetilde{n}_{ij}$ be such that

$$\underline{w} = \left(\frac{\widetilde{n}_{ij}}{n_j\left(\widetilde{n}_{ij}\right)}\right)^{\frac{1}{\eta}} \left(\frac{n_j\left(\widetilde{n}_{ij}\right)}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}$$

where

$$\max\left\{\underline{w}, \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}\right\} = \begin{cases} \underline{w} & \text{if } n_{ij} < \widetilde{n}_{ij} \\ \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} & \text{if } n_{ij} \geq \widetilde{n}_{ij}. \end{cases}$$

Note that the marginal revenue product is well defined and differentiable for all n_{ij}. Total labor costs are differentiable everywhere except at $n_{ij} = \widetilde{n}_{ij}$. However, the unconstrained labor supply curve is differentiable everywhere (note that n_j depends on n_{ij} and we suppress dependence on aggregates and competitor employment, both of which are taken as given),

$$\widehat{w}\left(n_{ij}\right) = \left(\frac{n_{ij}}{n_j\left(n_{ij}\right)}\right)^{\frac{1}{\eta}} \left(\frac{n_j\left(n_{ij}\right)}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}.$$

There are three possible first-order conditions, depending on the firm's optimal choice of n_{ij} relative to $\widetilde{n}_{ij}$,

$$mrpl\left(n_{ij}\right) = \begin{cases} \underline{w} & n_{ij} < \widetilde{n}_{ij} \\ \in \left[\underline{w}, \widehat{w}'\left(n_{ij}\right) n_{ij} + \widehat{w}\left(n_{ij}\right)\right] & n_{ij} = \widetilde{n}_{ij} \\ \widehat{w}'\left(n_{ij}\right) n_{ij} + \widehat{w}\left(n_{ij}\right) & n_{ij} > \widetilde{n}_{ij}. \end{cases}$$

Lemma 3 characterizes the firm's optimal choice of n_{ij}.

Lemma 3 - *Given competitor employment $\{n_{-ij}\}$, the firm's optimal choice of n_{ij} satisfies $mrpl\left(n_{ij}\right) \geq \underline{w}$.*

Proof: If $n_{ij} < \widetilde{n}_{ij}$ then $mrpl\left(n_{ij}\right) = \underline{w}$. If $n_{ij} = \widetilde{n}_{ij}$ then $mrpl\left(n_{ij}\right) \geq \underline{w}$ where we have used $\widehat{w}'\left(\widetilde{n}_{ij}\right)\widetilde{n}_{ij}+\widehat{w}\left(\widetilde{n}_{ij}\right)=\widehat{w}'\left(\widetilde{n}_{ij}\right)\widetilde{n}_{ij}+\underline{w}>\underline{w}$. If $n_{ij}>\widetilde{n}_{ij}$, we need to show that show that $\widehat{w}'\left(n_{ij}\right)n_{ij}+\widehat{w}\left(n_{ij}\right)$ is also increasing in n_{ij}, therefore $mrpl\left(n_{ij}\right)=\widehat{w}'\left(n_{ij}\right)n_{ij}+\widehat{w}\left(n_{ij}\right)>\widehat{w}'\left(\widetilde{n}_{ij}\right)\widetilde{n}_{ij}+\widehat{w}\left(\widetilde{n}_{ij}\right)>\underline{w}$. We can rewrite the marginal cost the firm as follows:

$$mc\left(n_{ij}\right)=w'\left(n_{ij}\right)n_{ij}+w\left(n_{ij}\right)=\left[\frac{w'\left(n_{ij}\right)n_{ij}}{w\left(n_{ij}\right)}+1\right]w\left(n_{ij}\right)=\left[\varepsilon^{Inv}\left(n_{ij}\right)+1\right]w\left(n_{ij}\right).$$

We can then show that $mc'(n_{ij})>0$ so long as $\varepsilon^{Inv\prime}(n_{ij})>0$:

$$mc'\left(n_{ij}\right)=\underbrace{\varepsilon^{Inv\prime}\left(n_{ij}\right)}_{\text{RHS positive if this is positive}}w\left(n_{ij}\right)+\left[\varepsilon^{Inv}\left(n_{ij}\right)+1\right]w'\left(n_{ij}\right)$$

Following the derivations in BHM this is true in the Cournot oligopsony problem of the firm since we have

$$\varepsilon^{Inv}\left(n_{ij}\right)=\frac{1}{\theta}s_{ij}+\left(1-s_{ij}\right)\frac{1}{\eta}$$

which is increasing (holding n_{-ij} fixed), as higher n_{ij} increases also w_{ij}, which increases s_{ij}, which pushes toward the larger $1/\theta$ term which is > $1/\eta$.■

Define $\overline{n}_{ij}$ by $mrpl\left(\overline{n}_{ij}\right)=\underline{w}$. Then by Lemma 3 we know that $mrpl\left(n_{ij}\right)\geq\underline{w}=mrpl\left(\overline{n}_{ij}\right)$, and hence $n_{ij}\leq\overline{n}_{ij}$. Therefore we can always set $\overline{n}_{ij}$ by $mrpl\left(\overline{n}_{ij}\right)=\underline{w}$ and this rationing constraint is non-binding away from the minimum wage, and weakly binding at the optimal value of employment when the firm is constrained by the minimum wage. Lemma 4 formally proves this result.

Lemma 4 - *It is (weakly) optimal for the firm to choose a rationing constraint* $\overline{n}_{ij}=\left(\alpha z_{ij}/\underline{w}\right)^{\frac{1}{1-\alpha}}$.

Proof: As above, define $\overline{n}_{ij}$ by $mrpl\left(\overline{n}_{ij}\right)=\underline{w}$. Note that $mrpl\left(n_{ij}\right)=\alpha z_{ij}n_{ij}^{\alpha-1}$, thus $\underline{w}=\alpha z_{ij}\overline{n}_{ij}^{\alpha-1}$, and $mrpl\left(\overline{n}_{ij}\right)=\underline{w}$. Also note that $mrpl\left(n\right)$ is *decreasing* in n. Conditional on competitor employment, we define three regions (I,II, and III) on the perceived inverse labor supply curve:

$$w^p\left(n_{ij},\overline{n}_{ij},n_j,N\right)=\begin{cases}\underbrace{\underline{w}}_{\text{Region III}} & \text{if } \underline{w}>\left(\frac{\overline{n}_{ij}}{n_j\left(\overline{n}_{ij}\right)}\right)^{\frac{1}{\eta}}\left(\frac{n_j\left(\overline{n}_{ij}\right)}{N}\right)^{\frac{1}{\theta}}N^{\frac{1}{\varphi}}\\ \max\left\{\underbrace{\underline{w}}_{\text{Region II}},\underbrace{\left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}}\left(\frac{n_j}{N}\right)^{\frac{1}{\theta}}N^{\frac{1}{\varphi}}}_{\text{Region I}}\right\} & \text{if } \underline{w}\leq\left(\frac{\overline{n}_{ij}}{n_j\left(\overline{n}_{ij}\right)}\right)^{\frac{1}{\eta}}\left(\frac{n_j\left(\overline{n}_{ij}\right)}{N}\right)^{\frac{1}{\theta}}N^{\frac{1}{\varphi}}.\end{cases}$$

Let **Region I** be the case that $\left(w_{ij},n_{ij},\overline{n}_{ij},n_j\right)$ are such that the firm is on the *sec-*

ond part of the second branch of $w\left(n_{ij},\overline{n}_{ij},n_j,N\right)$. Let **Region II** be the case that $\left(w_{ij},n_{ij},\overline{n}_{ij},n_j\right)$ are such that the firm is on the *first* part of the second branch of $w\left(n_{ij},\overline{n}_{ij},n_j,N\right)$. Let **Region III** be the case that $\left(w_{ij},n_{ij},\overline{n}_{ij},n_j\right)$ are such that the firm is on the first branch of $w\left(n_{ij},\overline{n}_{ij},n_j,N\right)$ (note that this *does not* require that $n_{ij}=\overline{n}_{ij}$, although this will be the case under firm optimality). We proceed by solving for the optimal n_{ij} in each Region and show that $\overline{n}_{ij}$ is weakly binding, and thus weakly optimal.

Region I. Suppose the firm is in Region I, then it is solving the problem (taking competitor employment as given),

$$\max_{n_{ij}\leq\overline{n}_{ij}} z_{ij}n_{ij}^{\alpha}-\widehat{w}\left(n_{ij}\right)n_{ij} \quad , \quad \widehat{w}\left(n_{ij}\right)=\left(\frac{n_{ij}}{n_j\left(n_{ij}\right)}\right)^{\frac{1}{\eta}}\left(\frac{n_j\left(n_{ij}\right)}{N}\right)^{\frac{1}{\theta}}N^{\frac{1}{\varphi}}$$

and hence has first order condition

$$\begin{aligned}
mrpl\left(n_{ij}^*\right) &= \widehat{w}'\left(n_{ij}^*\right)n_{ij}^*+\widehat{w}\left(n_{ij}^*\right) && , \text{ then since } \widehat{w}'\left(n_i\right)>0\\
&> \widehat{w}\left(n_{ij}^*\right) && , \text{ then since in Region I, then } \widehat{w}\left(n_{ij}^*\right)>\underline{w}\\
&> \underline{w}=mrpl(\overline{n}_{ij}) && , \text{ by the conjectured } \overline{n}_{ij},\ \underline{w}=mrpl(\overline{n}_{ij}).
\end{aligned}$$

Since $mrpl(n_{ij}^*)>mrpl(\overline{n}_{ij})$, and $mrpl$ is decreasing, $n_{ij}^*<\overline{n}_{ij}$. Therefore the constraint is slack. Note also that the value is independent of $\overline{n}_{ij}$.

Region II. Suppose the firm is in Region II, then $w_{ij}=\underline{w}$ and $\underline{w}\leq\left(\frac{\overline{n}_{ij}}{n_j\left(\overline{n}_{ij}\right)}\right)^{\frac{1}{\eta}}\left(\frac{n_j\left(\overline{n}_{ij}\right)}{N}\right)^{\frac{1}{\theta}}N^{\frac{1}{\varphi}}$. Define $\widetilde{n}_{ij}$ such that

$$\underline{w}=\left(\frac{\widetilde{n}_{ij}}{n_j\left(\widetilde{n}_{ij}\right)}\right)^{\frac{1}{\eta}}\left(\frac{n_j\left(\widetilde{n}_{ij}\right)}{N}\right)^{\frac{1}{\theta}}N^{\frac{1}{\varphi}}.$$

Note that for $n_{ij}\leq\widetilde{n}_{ij}$, then $\max\left\{\underline{w},\left(\frac{n_{ij}}{n_j\left(n_{ij}\right)}\right)^{\frac{1}{\eta}}\left(\frac{n_j\left(n_{ij}\right)}{N}\right)^{\frac{1}{\theta}}N^{\frac{1}{\varphi}}\right\}=\underline{w}$, and hence the firm is in Region II, while the firm is not in Region II for $n_{ij}>\widetilde{n}_{ij}$. Since the firm is in Region II, then also have $\underline{w}\leq\left(\frac{\overline{n}_{ij}}{n_j\left(\overline{n}_{ij}\right)}\right)^{\frac{1}{\eta}}\left(\frac{n_j\left(\overline{n}_{ij}\right)}{N}\right)^{\frac{1}{\theta}}N^{\frac{1}{\varphi}}$, which by monotonicity of the labor supply curve implies that $\widetilde{n}_{ij}\leq\overline{n}_{ij}$. Therefore the n_{ij} for which the firm is in Region II are all weakly less than $\overline{n}_{ij}$. Note that this *does not* require knowing anything about the $mrpl_{ij}$, its simply by definition of Region II. Note also that the value is independent of $\overline{n}_{ij}$.

Region III. Suppose the firm is in Region III, then $w_{ij} = \underline{w}$ for all $n_{ij} \leq \overline{n}_{ij}$. Therefore the firm is solving:

$$\max_{n_{ij}} \; z_{ij} n_{ij}^{\alpha} - \underline{w} n_{ij}$$

and hence has the first order condition $mrpl\left(n_{ij}^{*}\right) = \underline{w} = mrpl\left(\overline{n}_{ij}\right)$ therefore the constraint is weakly binding.■

Applying Lemma 4, we can write the firm problem with the constraint $\overline{n}_{ij} = \left(\frac{\alpha z_{ij}}{\underline{w}}\right)^{\frac{1}{1-\alpha}}$ imposed:

$$\max_{n_{ij}} z_{ij} n_{ij}^{\alpha} - w\left(n_{ij}, \overline{n}_{ij}, n_j, N\right) n_{ij}$$

subject to $n_{ij} \leq \overline{n}_{ij}$ and $\overline{n}_{ij} = \left(\frac{\alpha z_{ij}}{\underline{w}}\right)^{\frac{1}{1-\alpha}}$ and the *perceived labor supply curve*

$$w^{p}\left(n_{ij}, \overline{n}_{ij}, n_j, N\right) = \begin{cases} \underline{w} & \text{if } \underline{w} > \left(\frac{\overline{n}_{ij}}{n_j\left(\overline{n}_{ij}\right)}\right)^{\frac{1}{\eta}} \left(\frac{n_j\left(\overline{n}_{ij}\right)}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} \\ \max\left\{\underline{w}, \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}\right\} & \text{if } \underline{w} \leq \left(\frac{\overline{n}_{ij}}{n_j\left(\overline{n}_{ij}\right)}\right)^{\frac{1}{\eta}} \left(\frac{n_j\left(\overline{n}_{ij}\right)}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} \end{cases}$$

This is the problem described in the main text.

SUPPLEMENTAL APPENDIX TO

Minimum Wages, Efficiency and Welfare

David Berger, Kyle Herkenhoff, Simon Mongey

April 24, 2024

The indexing of this ***Supplemental Appendix*** follows on from the ***Online Appendix***. Section E contains additional details on calibration of preference parameters. Section F provides a pedagogical step-by-step solution of a simplified version of the model and the algorithm for solving the minimum wage economy. Section G contains mathematical derivations for the full quantitative model with household heterogeneity. Section I contains derivations of the solution of the model under the tax and transfer system in Section 6.

E Disciplining preference parameters

This Section details how we use recent evidence from Golosov, Graber, Mogstad, and Novgorodsky (2021) to discipline preference parameters σ and φ.

Background. Consider a budget constraint, where b_i is unearned income and $\mathcal{T}$ gives taxes and transfers which depend on pre-tax labor income y_i:

$$c_i = y_i - \mathcal{T}(y_i) + b_i$$

Totally differentiating with respect to b_i:

$$\frac{dc_i}{db_i} = \frac{dy_i}{db_i} - \frac{d\mathcal{T}_i}{db_i} + 1 \quad \text{, which we can write} \quad MPC_i = MPE_i - MPT_i + 1$$

Table 4.1 of Golosov, Graber, Mogstad, and Novgorodsky (2021, henceforth GGMN) gives estimates of the marginal propensity to consume (MPC) and marginal propensity to earn (MPE) for different income groups, where lottery winnings are used as an instrument for the endogenous variable b_i. For example, results are of the type: *An extra dollar in unearned income leads to a $MPE = -0.52$ cent reduction in labor earnings.* We show how their results can be used to discipline preference parameters (φ, σ) in a simple labor supply setting that is consistent with our model.

Derivation. Consider the following individual problem, where preferences are as in the main text, and $y = wn$, where w is taken as given:

$$u(c,n) = \frac{c^{1-\sigma}}{1-\sigma} - \frac{1}{\overline{\varphi}^{1/\varphi}} \frac{n^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} \quad \text{subject to} \quad c = wn + \mathcal{T}(wn) + b \tag{E1}$$

Optimality conditions for c and n give labor supply, which can be expressed in terms of earnings:

$$y = \overline{\varphi} c^{-\varphi\sigma} w^{\varphi+1} \left(1 - \mathcal{T}'(y)\right)^{\varphi}$$

Totally differentiating with respect to b

$$\frac{dy}{db} = -\varphi\sigma \frac{dc}{db}\left(\frac{y}{c}\right) - \varphi\left(\frac{\mathcal{T}''(y)\,y}{1-\mathcal{T}'(y)}\right)\frac{dy}{db}.$$

Now suppose that post-tax labor earnings were of the form used in Heathcote, Storesletten, and Violante (2020, henceforth HSV): $y - \mathcal{T}(y) = \lambda y^{1-\tau}$. In this case, the elasticity term is simply the progressivity of taxes, τ.

$$\frac{dy}{db} = -\varphi\sigma \frac{dc}{db}\left(\frac{y}{c}\right) - \varphi\tau\frac{dy}{db}.$$

Using the definitions of MPC, MPE, the average propensity to consume $APC = c/y$, and after rearranging, we have a closed-form relationship between σ and φ, given data on $\{MPC, MPE, APC, \tau\}$:

$$\varphi = -\frac{1}{\sigma\frac{MPC}{MPE}\frac{1}{APC} - \tau}. \tag{E2}$$

If we let $\sigma = 1$ and $\tau = 0$, it is straightforward to observe that a lower MPC and higher MPE in absolute terms (as will be the case for richer households), requires a higher φ.

$$\varphi = \frac{|MPE|}{MPC} APC.$$

Data. We use BLS data to compute APC for non-high-school, high-school, and college completion households. We map these into the four quartiles of income groups in GGMN Table 4.1 as given in the following table. We take a value of $\tau = 0.181$ from HSV.

	All	**Group**		
BLS category GGMN category		Non-High School Q1	High school Q2-Q3	Completed college Q4
APC (BLS)	0.69	0.73	0.71	0.67
MPE (GGMN)	-0.5227	-0.3080	-0.5549	-0.6735
MPC (GGMN)	0.5836	0.7315	0.5429	0.4990

Table E1: Data used in calibrating preference parameters

Results. Using equation (E2), we can then determine φ given σ. Figure E1 plots $\varphi(\sigma)$ for $\sigma \in [1, 2]$. As a benchmark, with log preferences, and when calibrated to the whole sample values, $\varphi(1) = 0.65$. For low income (Q1) households $\varphi(1) = 0.32$, for high

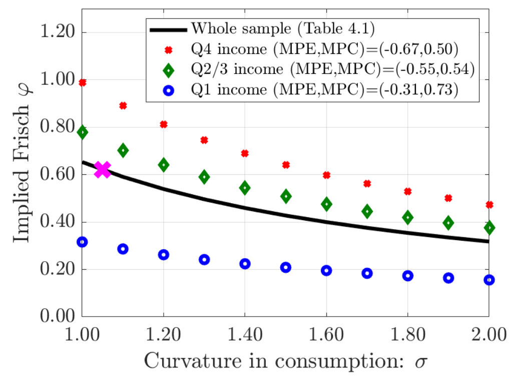


Figure E1: Implied parameters

Notes: Given a value for the coefficient of relative risk aversion σ, this figure plots the Frisch elasticity of labor supply φ required for the optimality conditions of the simple labor supply model E1 to be consistent with (i) empirical measures of the marginal propensity to earn and marginal propensity to consume following changes in unearned income from Golosov, Graber, Mogstad, and Novgorodsky (2021), (ii) estimates of the average propensity to consume from the BLS, (iii) estimates of the progressivity of post-tax labor income to pre-tax-and-transfer income from Heathcote, Storesletten, and Violante (2020).

income households $\varphi(1) = 0.987$. High income (Q4) households have higher MPE's, and their MPC is lower, reducing $|MPC/MPE|$, and requiring a higher φ. The pink cross corresponds to $(\sigma, \varphi) = (1.05, 0.62)$, which are the values used in the baseline calibration of our model (see Table 1).

F Pedagogical example & algorithm

The aim of this section is to clearly lay out the algorithm for solving the minimum wage equilibrium, and to present a full solution of a simplified model, which may be pedagogically useful relative to the extensive derivations in Appendix D. The algorithm for the minimum wage equilibrium is nested in the broader solution to the equilibrium of the model described in Appendix G.

For ease of exposition, we lay out the minimum wage problem (i) ignoring capital, (ii) consider an economy with a single type of household, (iii) to simplify exposition we also consider GHH preferences, which are not used in the main text, (iv) as well as a static environment, (v) set the coefficient on labor in utility $\overline{\varphi} = 1$. We derive conditions for this simplified economy and then present the algorithm.

F.1 Pedagogical example

Consider the household problem with the rationing constraint $n_{ij} \leq \overline{n}_{ij}$. For ease of interpretation we attach multiplier $\zeta_{ij} = \lambda w_{ij} \left(1 - p_{ij}\right)$ to the rationing constraint, normalized by the household budget multiplier λ:

$$U_0 = \max_{\{n_{ij},c_{ij}\}} u\left(C - \frac{N^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}\right)$$

$$\begin{aligned}
C &= \int \sum_{i\in j} w_{ij} n_{ij}\, dj + \Pi \qquad [\lambda] \\
n_{ij} &\leq \overline{n}_{ij} \qquad [\lambda w_{ij}\left(1-p_{ij}\right)] \\
C &= \int \sum_{i\in j} c_{ij} dj \quad , \quad N = \left[\int n_j^{\frac{\theta+1}{\theta}}\, dj\right]^{\frac{\theta}{\theta+1}} \quad , \quad n_j = \left[\sum_{i\in j} n_j^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}}
\end{aligned}$$

The first order condition for n_{ij} yields

$$\begin{aligned}
\lambda w_{ij} - \lambda w_{ij}\left(1-p_{ij}\right) &= u'\left(\cdot\right)\left(\frac{\partial n_j}{\partial n_{ij}}\right)\left(\frac{\partial N}{\partial n_j}\right) N^{\frac{1}{\varphi}} \\
\lambda w_{ij} p_{ij} &= u'\left(\cdot\right)\left(\frac{\partial n_j}{\partial n_{ij}}\right)\left(\frac{\partial N}{\partial n_j}\right) N^{\frac{1}{\varphi}}
\end{aligned}$$

The first order condition for consumption yields $u'\left(\cdot\right) = \lambda$. Define the *shadow wage* $\widetilde{w}_{ij} = p_{ij} w_{ij}$. Use the first order condition for consumption $u'\left(\cdot\right) = \lambda$, and use the derivatives of N and n_j:

$$\widetilde{w}_{ij} = \left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} \qquad (*)$$

Now define the *shadow wage indexes*

$$\widetilde{w}_j = \left[\sum_{i\in j} \widetilde{w}_{ij}^{1+\eta}\right]^{\frac{1}{1+\eta}} \quad , \quad \widetilde{W} = \left[\int \widetilde{w}_j^{1+\theta} dj\right]^{\frac{1}{1+\theta}} .$$

Using these definitions in $(*)$ along with the definition of n_j:

$$\begin{aligned}
\sum_{i\in j} \widetilde{w}_{ij}^{1+\eta} &= \left[\left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}\right]^{1+\eta} \sum_{i\in j} \left(\frac{n_{ij}}{n_j}\right)^{\frac{1+\eta}{\eta}} \\
\widetilde{w}_j &= \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}
\end{aligned}$$

Using this along with the definition of N:

$$\begin{aligned}
\int \widetilde{w}_j^{1+\theta} dj &= \left[N^{\frac{1}{\varphi}}\right]^{1+\theta} \int \left(\frac{n_j}{N}\right)^{\frac{1+\theta}{\theta}} dj \\
\widetilde{W} &= N^{\frac{1}{\varphi}}
\end{aligned}$$

Note that $\widetilde{W}N \neq \int \sum_{i\in j} w_{ij}n_{ij}dj$, however the aggregate labor supply $N = \widetilde{W}^{\varphi}$ is *as if*, the household had maximized

$$U_0 = \max_{C,N} u\left(C - \frac{N^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}\right) \quad \text{subject to} \quad C = \widetilde{W}N + \Pi.$$

This makes clear the extent to which the shadow wage index $\widetilde{W}$ captures the full distribution of binding minimum wages. Note that shadow wages aggregate:

$$\begin{aligned}
\widetilde{w}_{ij}n_{ij} &= n_{ij}^{\frac{1+\eta}{\eta}}\left(\frac{1}{n_j}\right)^{\frac{1}{\eta}}\left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} \\
\sum_{i\in j}\widetilde{w}_{ij}n_{ij} &= \left[\sum_{i\in j} n_{ij}^{\frac{1+\eta}{\eta}}\right]\left(\frac{1}{n_j}\right)^{\frac{1}{\eta}}\left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}} \\
\sum_{i\in j}\widetilde{w}_{ij}n_{ij} &= n_j\widetilde{w}_j
\end{aligned}$$

Shadow shares - We can define the shadow share $\widetilde{s}_{ij}$ as

$$\widetilde{s}_{ij} := \frac{\widetilde{w}_{ij}n_{ij}}{\sum_{i\in j}\widetilde{w}_{ij}n_{ij}}.$$

Substituting in the labor supply system $(*)$ for $\widetilde{w}_{ij}$

$$\widetilde{s}_{ij} := \frac{n_{ij}^{\frac{1+\eta}{\eta}}}{\sum_{i\in j} n_{ij}^{\frac{1+\eta}{\eta}}} = \left(\frac{n_{ij}}{n_j}\right)^{\frac{1+\eta}{\eta}} = \left(\frac{\widetilde{w}_{ij}}{\widetilde{w}_j}\right)^{1+\eta}$$

The firm's problem is

$$\pi_{ij} = \max_{n_{ij}} z_{ij}n_{ij}^{\alpha} - w_{ij}n_{ij}$$

subject to

$$\begin{aligned}
n_{ij} &= \left(\frac{\widetilde{w}_{ij}}{\widetilde{w}_j}\right)^{\eta}\left(\frac{\widetilde{w}_j}{\widetilde{W}}\right)^{\theta} N \\
w_{ij} &\geq \underline{w}
\end{aligned}$$

Let $r_{ij} \in \{1,2,3\}$ denote the region that the firm is in.

Region I - If the firm is in Region I, then its wage is the optimal markdown on the marginal revenue product of labor

$$w_{ij} = \mu_{ij}\alpha z_{ij}n_{ij}^{\alpha-1} \quad , \quad p_{ij} = 1 \quad , \quad \widetilde{w}_{ij} = w_{ij} \quad , \quad n_{ij} = \left(\frac{w_{ij}}{\widetilde{w}_j}\right)^{\eta}\left(\frac{\widetilde{w}_j}{\widetilde{W}}\right)^{\theta}\widetilde{W}^{\varphi}$$

where the markdown depends on its shadow share of the labor market. That is, $\mu_{ij} =$

$\mu\left(\widetilde{s}_{ij}\right)$, where $\mu\left(\widetilde{s}_{ij}\right)=\frac{\varepsilon\left(\widetilde{s}_{ij}\right)}{\varepsilon\left(\widetilde{s}_{ij}\right)+1}$. We have shown that

$$\widetilde{s}_{ij}=\left(\frac{\widetilde{w}_{ij}}{\widetilde{w}_j}\right)^{1+\eta} \qquad \Longrightarrow \widetilde{w}_j=\widetilde{w}_{ij}\widetilde{s}_{ij}^{-\frac{1}{1+\eta}}$$

Using these, we can write:

$$w_{ij}=\left[\mu\left(\widetilde{s}_{ij}\right)\alpha z_{ij}\widetilde{s}_{ij}^{-\frac{(1-\alpha)(\eta-\theta)}{1+\eta}}\widetilde{W}^{(1-\alpha)(\theta-\varphi)}\right]^{\frac{1}{1+\theta(1-\alpha)}}$$

Region II - In Region II, then

$$w_{ij}=\underline{w} \quad , \quad p_{ij}=1 \quad , \quad \widetilde{w}_{ij}=\underline{w} \quad , \quad n_{ij}=\left(\frac{\underline{w}}{\widetilde{w}_j}\right)^{\eta}\left(\frac{\widetilde{w}_j}{\widetilde{W}}\right)^{\theta}N$$

Region III - In Region III, then

$$w_{ij}=\alpha z_{ij}n_{ij}^{\alpha-1} \quad , \quad p_{ij}<1 \quad , \quad \widetilde{w}_{ij}=p_{ij}\underline{w} \quad , \quad n_{ij}=\left(\frac{p_{ij}\underline{w}}{\widetilde{w}_j}\right)^{\eta}\left(\frac{\widetilde{w}_j}{\widetilde{W}}\right)^{\theta}N$$

F.2 Minimum wage solution algorithm

We implement the following solution algorithm. We denote the *Region* that a firm is in by $r_{ijt}\in\{I,II,III\}$. Initialize the algorithm by (i) guessing a value for $\widetilde{W}^{(0)}$, (ii) assuming all firms are in *Region I*, $r_{ij}^{(0)}=I$, which implies guessing $p_{ij}^{(0)}=1$. These will all be updated in the algorithm.

1. Solve all market equilibria in shadow shares.

1. Guess shadow shares $\widetilde{s}_{ij}^{(0)}$.
2. *Region I* - Using the above optimality condition

$$w_{ij}=\left[\mu\left(\widetilde{s}_{ij}\right)\alpha z_{ij}\widetilde{s}_{ij}^{(0)-\frac{(1-\alpha)(\eta-\theta)}{1+\eta}}\widetilde{W}^{(0)(1-\alpha)(\theta-\varphi)}\right]^{\frac{1}{1+\theta(1-\alpha)}}$$

3. *Regions II, III* - Here the minimum wage is binding so set $w_{ij}=\underline{w}$.
4. Given the guess $p_{ij}^{(k)}$ and w_{ij}, compute the shadow wage: $\widetilde{w}_{ij}=p_{ij}w_{ij}$.
5. With all shadow wages in hand, update shadow shares using $\widetilde{w}_{ijt}$:

$$\widetilde{s}_{ij}^{(l+1)}=\frac{\widetilde{w}_{ij}^{1+\eta}}{\sum_{i\in j}\widetilde{w}_{ij}^{1+\eta}}.$$

6. Iterate over (b)-(e) until shadow shares converge: $\widetilde{s}_{ij}^{(l+1)}=\widetilde{s}_{ij}^{(l)}$.

2. Recover employment. Here we use the wages from the previous step plus the current guess of each firms' region. First aggregate $\widetilde{w}_{ij}$ to compute $\widetilde{w}_j$ and $\widetilde{W}$. Then by region $r^{(k)}_{ijt}$:

1. *Region I* - Firm is unconstrained:

$$n_{ij} = \left(\frac{w_{ij}}{\widetilde{w}_j}\right)^{\eta} \left(\frac{\widetilde{w}_j}{\widetilde{W}}\right)^{\theta} \widetilde{W}^{\varphi}$$

2. *Region II* - Firm is constrained and n_{ij} determined by household labor supply curve at $\underline{w}$:

$$n_{ij} = \left(\frac{\underline{w}}{\widetilde{w}_j}\right)^{\eta} \left(\frac{\widetilde{w}_j}{\widetilde{W}}\right)^{\theta} \widetilde{W}^{\varphi}$$

3. *Region III* - Firm is constrained and n_{ijt} determined by firm labor demand curve at $\underline{w}$:

$$\underline{w} = \alpha z_{ij} n_{ij}^{\alpha-1} \quad \Longrightarrow \quad n_{ij} = \left(\frac{\alpha z_{ij}}{\underline{w}}\right)^{\frac{1}{1-\alpha}}.$$

3. Update the multipliers: $p^{(k)}_{ij}$.

1. Aggregate n_{ij} to compute n_j and N.
2. Update p_{ij} from the *household's* first order conditions: $\widetilde{w}_{ij} = p_{ij} w_{ij}$

$$p^{(k+1)}_{ij} = \frac{\left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{\theta}} N^{\frac{1}{\varphi}}}{w_{ij}}$$

4. Update $\widetilde{W}^{(k)}$.

1. Compute $\widetilde{w}_{ij} = p^{(k+1)}_{ij} w_{ij}$
2. Use $\widetilde{w}_{ij}$ to update the aggregate shadow wage index to $\widetilde{W}^{(k+1)}$.

5. Update firm regions. For each region:.

1. Compute the marginal product of labor of all firms $mrpl_{ij} = \alpha z_{ij} n_{ij}^{\alpha-1}$.
2. If in market j there exists a firm in *Region I* with $w_{ij} < \underline{w}$, then move the firm with the lowest wage into *Region II*
3. If in market j there exists a firm that was initially in *Region II* and has a marginal product of labor that is less than marginal cost ($\underline{w}$), move that firm into *Region III*

Iterate over (1) to (5) until $p^{(k+1)}_{ij} = p^{(k)}_{ij}$ and $\widetilde{W}^{(k+1)} = \widetilde{W}^{(k)}$ and $r^{(k+1)}_{ij} = r^{(k)}_{ij}$.

G Mathematical details - Full quantitative model

- We first derive results for the competitive equilibrium, then the government's allocation problem. We then use results from the competitive equilibrium to prove that the solution to the government's allocation problem can be decentralized in a competitive equilibrium with revenue neutral lump sum taxes

G.1 Competitive equilibrium

G.1.1 Household problem - Labor supply system, shadow wages

- In the competitive equilibrium, household h solves the following problem:

$$\max_{c_{ht},n_{ht}} \sum_{t=0}^{\infty} \beta^t \left[\frac{\left(c_{ht}/\pi_h\right)^{1-\sigma}}{1-\sigma} - \frac{1}{\widetilde{\varphi}_h^{1/\varphi}} \frac{n_{ht}^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}} \right]$$

where $\widetilde{\varphi}_h = \overline{\varphi}_h \pi_h^{1+\varphi}$ is adjusted for the measure of workers of the household,

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

subject to the budget constraint

$$c_{ht} + k_{ht+1} = \int \sum_{i \in j} w_{ijht} n_{ijht} dj + R_t k_{ht} + (1-\delta) k_{ht} + \kappa_h \Pi_t.$$

with the initial condition $k_{h0} = \kappa_h K_0$.

- Since we focus on steady-state we normalize the price of consumption to one.
- In the text we refer to these preferences as $u^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)$:

$$u^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right) = \frac{\left(c_{ht}/\pi_h\right)^{1-\sigma}}{1-\sigma} - \frac{1}{\widetilde{\varphi}_h^{1/\varphi}} \frac{n_{ht}^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}$$

- The household is also subject to the firm by firm rationing constraints: $n_{ijht} \leq \overline{n}_{ijht}$.
- Let $\beta^t \nu_{ht}$ be the multiplier on the household's budget constraint and write the multiplier on the rationing constraint as $\zeta_{ijht} = \beta^t \nu_{ht} w_{ijht} \left(1 - p_{ijht}\right)$.

- The the household's Lagrangean features the following terms in n_{ijht}

$$\mathcal{L} = \cdots + \beta^t u^h \left(\frac{c_{ht}}{\pi_h}, n_{ht} \right) + \cdots + \beta^t \nu_{ht} w_{ijht} n_{ijht} + \beta^t \nu_{ht} w_{ijht} \left(1 - p_{ijht}\right) \left[\overline{n}_{ijht} - n_{ijht} \right] + \ldots$$

$$\mathcal{L} = \cdots + u^h \left(\frac{c_{ht}}{\pi_h}, n_{ht} \right) + \cdots + \beta^t \nu_{ht} \left\{ w_{ijht} p_{ijht} \right\} n_{ijht} + \beta^t \nu_{ht} w_{ijht} \left(1 - p_{ijht}\right) \overline{n}_{ijht} + \ldots$$

- The first order condition for consumption is

$$u_c^h \left(\frac{c_{ht}}{\pi_h}, n_{ht} \right) = \nu_{ht}$$

- The first order condition for labor supply is

$$\begin{aligned} \nu_{ht} w_{ijht} p_{ijht} &= -u_n^h \left(\frac{c_{ht}}{\pi_h}, n_{ht} \right) \frac{\partial n_{ht}}{\partial n_{jht}} \frac{\partial n_{jht}}{\partial n_{ijht}} \\ w_{ijht} p_{ijht} &= -\frac{u_n^h \left(\frac{c_{ht}}{\pi_h}, n_{ht} \right)}{u_c^h \left(\frac{c_{ht}}{\pi_h}, n_{ht} \right)} \left(\frac{n_{jht}}{n_{ht}} \right)^{\frac{1}{\theta}} \left(\frac{n_{ijht}}{n_{jht}} \right)^{\frac{1}{\eta}} \end{aligned}$$

- Define the *shadow wage* by $\widetilde{w}_{ijht} := w_{ijht} p_{ijht}$.

- Then

$$\widetilde{w}_{ijht} = -\frac{u_n^h \left(\frac{c_{ht}}{\pi_h}, n_{ht} \right)}{u_c^h \left(\frac{c_{ht}}{\pi_h}, n_{ht} \right)} \left(\frac{n_{jht}}{n_{ht}} \right)^{\frac{1}{\theta}} \left(\frac{n_{ijht}}{n_{jht}} \right)^{\frac{1}{\eta}}.$$

- Now define the following *shadow wage indexes*:

$$\widetilde{w}_{jht} = \left[\sum_{i \in j} \widetilde{w}_{ijht}^{1+\eta} \right]^{\frac{1}{1+\eta}} \quad , \quad \widetilde{w}_{ht} = \left[\int \widetilde{w}_{jht}^{1+\theta} dj \right]^{\frac{1}{1+\eta}}$$

- Using this

$$\sum_{i\in j}\widetilde{w}_{ijht}^{1+\eta} = \left[-\frac{u_n^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)}{u_c^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)}\left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1}{\theta}}\right]^{1+\eta}\sum_{i\in j}\left(\frac{n_{ijht}}{n_{jht}}\right)^{\frac{1+\eta}{\eta}}$$

$$\widetilde{w}_{jht} = -\frac{u_n^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)}{u_c^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)}\left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1}{\theta}}$$

$$\widetilde{w}_{jht}^{1+\theta} = \left[-\frac{u_n^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)}{u_c^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)}\right]\left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1+\theta}{\theta}}$$

$$\int \widetilde{w}_{jht}^{1+\theta}dj = \left[-\frac{u_n^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)}{u_c^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)}\right]\int\left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1+\theta}{\theta}}dj$$

$$\widetilde{w}_{ht} = -\frac{u_n^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)}{u_c^h\left(\frac{c_{ht}}{\pi_h}, n_{ht}\right)}$$

- Using our form of preferences, this gives the household h labor supply curve:

$$n_{ht} = \overline{\varphi}_h\pi_h\widetilde{w}_{ht}^{\varphi}\left(\frac{c_{ht}}{\pi_h}\right)^{-\varphi\sigma}$$

- Using this we can show that shadow wages aggregate, as claimed in the text,

- First across markets:

$$\widetilde{w}_{ijht} = \widetilde{w}_{ht}\left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1}{\theta}}\left(\frac{n_{ijht}}{n_{jht}}\right)^{\frac{1}{\eta}}.$$

$$\widetilde{w}_{ijht}^{1+\eta} = \left[\widetilde{w}_{ht}\left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1}{\theta}}\right]^{1+\eta}\left(\frac{n_{ijht}}{n_{jht}}\right)^{\frac{1+\eta}{\eta}}$$

$$\left[\sum_{i\in j}\widetilde{w}_{ijht}^{1+\eta}\right]^{\frac{1}{1+\eta}} = \widetilde{w}_{ht}\left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1}{\theta}}$$

$$\widetilde{w}_{jht} = \widetilde{w}_{ht}\left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1}{\theta}}$$

$$\widetilde{w}_{jht}n_{jht} = \widetilde{w}_{ht}n_{jht}\times\left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1+\theta}{\theta}}$$

$$\int\widetilde{w}_{jht}n_{jht}dj = \widetilde{w}_{ht}n_{jht}$$

- Then using these results, across firms within a market:

$$\begin{aligned}
\widetilde{w}_{ijht} &= \widetilde{w}_{ht}\left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1}{\theta}}\left(\frac{n_{ijht}}{n_{jht}}\right)^{\frac{1}{\eta}} \\
\widetilde{w}_{ijht} &= \widetilde{w}_{jht}\left(\frac{n_{ijht}}{n_{jht}}\right)^{\frac{1}{\eta}} \\
\widetilde{w}_{ijht}n_{ijht} &= \widetilde{w}_{jht}n_{jht}\times\left(\frac{n_{ijht}}{n_{jht}}\right)^{\frac{1+\eta}{\eta}} \\
\sum_{i\in j}\widetilde{w}_{ijht}n_{ijht} &= \widetilde{w}_{jht}n_{jht}
\end{aligned}$$

- Summarizing results so far, we have:

$$\begin{aligned}
\widetilde{w}_{ijht} &= \left(\frac{n_{ijht}}{n_{jht}}\right)^{\frac{1}{\eta}}\widetilde{w}_{jht} \\
\widetilde{w}_{jht} &= \left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1}{\theta}}\widetilde{w}_{ht} \\
\widetilde{w}_{ht}n_{jht} &= \int\widetilde{w}_{jht}n_{jht}dj \\
\widetilde{w}_{jht}n_{jht} &= \sum_{i\in j}\widetilde{w}_{ijht}n_{ijht}
\end{aligned}$$

- Note that these can be combined to give the entire labor supply system of household h in shadow wages:

$$\begin{aligned}
n_{ijht} &= \left(\frac{\widetilde{w}_{ijht}}{\widetilde{w}_{jht}}\right)^{\eta}\left(\frac{\widetilde{w}_{jht}}{\widetilde{w}_{ht}}\right)^{\theta}n_{ht} \\
n_{ht} &= \overline{\varphi}_h\pi_h^{1+\varphi\sigma}\widetilde{w}_{ht}^{\varphi}c_{ht}^{-\varphi\sigma}
\end{aligned}$$

- A hey result, used below, is that if the household received lump sum transfers T_h, then the same labor supply system would be obtained.
- Now consider our results regarding shadow shares. We define the *shadow share* as

$$\widetilde{s}_{ijht} := \frac{\widetilde{w}_{ijht}n_{ijht}}{\sum_{i\in j}\widetilde{w}_{ijht}n_{ijht}}.$$

- Using the above aggregation results, labor supply system, and definition of the aggregator n_{jht}:

$$\widetilde{s}_{ijht}=\frac{\widetilde{w}_{ijht}n_{ijht}}{\widetilde{w}_{jht}n_{jht}}=\left(\frac{\widetilde{w}_{ijht}}{\widetilde{w}_{jht}}\right)^{1+\eta}=\left(\frac{n_{ijht}}{n_{jht}}\right)^{\frac{1+\eta}{\eta}}=\frac{\partial\log n_{ijht}}{\partial\log n_{jht}}$$

which we use below in the firm optimality conditions.

G.1.2 Firm optimality

- **Simplifying the firm problem -** First we simplify the firm problem by separating it out across types and optimizing out capital for each type of worker:
- Consider the maximization problem of the firm in the text:

$$\pi_{ij}=\max_{\left\{n_{ijh},h_{ijh}\right\}_{h=1}^{h}}\overline{Z}z_{ij}\sum_{h=1}^{H}\left(\left[\xi_k n_{ijk}\right]^{\gamma}k_{ijk}^{1-\gamma}\right)^{\alpha}-R\sum_{h=1}^{H}k_{ijk}-\sum_{h=1}^{H}w_{ijh}n_{ijh}$$

  subject to the labor supply system and minimum wage constraints.

- First observe that this can separated out by type of worker h.
- The problem for type h labor at the firm is

$$\pi_{ijh}=\max_{n_{ijh},k_{ijh}}\overline{Z}z_{ij}\left(\left[\xi_h n_{ijh}\right]^{\gamma}k_{ijh}^{1-\gamma}\right)^{\alpha}-Rk_{ijh}-w_{ijh}n_{ijh}$$

- We first optimize out capital. This yields the objective function

$$\pi_{ijh}=\max_{n_{ijh}}\widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}}-w_{ijh}n_{ijh}$$

  where

$$\begin{aligned}\widetilde{Z}&=\overline{Z}^{\frac{1}{1-(1-\gamma)\alpha}}\\ \widetilde{z}_{ij}&=\left[1-(1-\gamma)\,\alpha\right]\left(\frac{(1-\gamma)\,\alpha}{R}\right)^{\frac{(1-\gamma)\alpha}{1-(1-\gamma)\alpha}}z_{ij}^{\frac{1}{1-(1-\gamma)\alpha}}\\ \widetilde{\xi}_h&=\xi_h^{\widetilde{\alpha}}\\ \widetilde{\alpha}&=\frac{\gamma\alpha}{1-(1-\gamma)\,\alpha}\end{aligned}$$

- We denote output net of capital expenses as $\widetilde{y}_{ijh}:=\widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}}$.

- We can also define a market-level aggregate $\widetilde{y}_{jh} = \sum_{i\in j} \widetilde{y}_{ijh}$, and a type-level aggregate $\widetilde{y}_h = \int \widetilde{y}_{jh} dj$.
- Note that

$$y_{ijh} = \frac{\widetilde{y}_{ijh}}{1-(1-\gamma)\alpha} \quad , \quad y_{jh} = \frac{\widetilde{y}_{jh}}{1-(1-\gamma)\alpha} \quad , \quad y_h = \frac{\widetilde{y}_h}{1-(1-\gamma)\alpha}.$$

- Using the simplified problem we now consider optimality of the firm in each of the three regions described in the text.
- **Region I - Unconstrained**
  - Consider an unconstrained firm. Its problem is

$$\pi_{ijh} = \max_{n_{ijh}} \widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}} - w_{ijh}n_{ijh}$$

subject to its wage being given by the above labor supply system:

$$w\left(n_{ijht}\right) = \left(\frac{n_{ijht}}{n_{jht}}\right)^{\frac{1}{\eta}} \left(\frac{n_{jht}}{n_{ht}}\right)^{\frac{1}{\theta}} \widetilde{w}_{ht}.$$

  - The first order condition is

$$\begin{aligned}
w_{ijh} + w'\left(n_{ijh}\right)n_{ijh} &= \widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}-1} \\
w_{ijh}\left(1 + \frac{w'\left(n_{ijh}\right)n_{ijh}}{w_{ijh}}\right) &= \widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}-1} \\
w_{ijh}\left(1 + \frac{1}{\varepsilon_{ijh}}\right) &= \widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}-1} \\
w_{ijh} &= \frac{\varepsilon_{ijh}}{1+\varepsilon_{ijh}}\widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}-1}
\end{aligned}$$

where using the inverse labor supply curve gives

$$\begin{aligned}
\frac{1}{\varepsilon_{ijh}} &:= \frac{w'\left(n_{ijh}\right)n_{ijh}}{w_{ijh}} = \frac{\partial \log w_{ijh}}{\partial \log n_{ijh}} = \frac{1}{\eta} + \left(\frac{1}{\theta} - \frac{1}{\eta}\right)\frac{\partial \log n_{jh}}{\partial \log n_{ijh}} = \frac{1}{\eta} + \left(\frac{1}{\theta} - \frac{1}{\eta}\right)\widetilde{s}_{ijh} \\
\varepsilon_{ijh} &= \left[\frac{1}{\eta} + \left(\frac{1}{\theta} - \frac{1}{\eta}\right)\widetilde{s}_{ijh}\right]^{-1}.
\end{aligned}$$

– Therefore

$$w_{ijh} = \mu_{ijh} \widetilde{\alpha} \widetilde{\widetilde{Z}} \widetilde{z}_{ij} \widetilde{\widetilde{\xi}}_h n_{ijh}^{\widetilde{\alpha}-1}$$

where the markdown depends on the firms' elasticity of labor supply.

– Note that since $p_{ijh} = 1$ since the firm is unconstrained, then $\widetilde{w}_{ijh} = p_{ijh} w_{ijh} = w_{ijh}$, so

$$\widetilde{w}_{ijh} = \mu_{ijh} \times \widetilde{\alpha} \widetilde{\widetilde{Z}} \widetilde{z}_{ij} \widetilde{\widetilde{\xi}}_h n_{ijh}^{\widetilde{\alpha}-1}$$

- **Region III - Constrained, on labor demand curve**

  – Now consider a constrained firm in Region III, this firm's problem is

$$\pi_{ijh} = \max_{n_{ijh}} \widetilde{\widetilde{Z}} \widetilde{z}_{ij} \widetilde{\widetilde{\xi}}_h n_{ijh}^{\widetilde{\alpha}} - \underline{w} n_{ijh}$$

  – The solution to this problem is to choose employment to equate the marginal revenue product of labor to the minimum wage:

$$\underline{w} = \widetilde{\alpha} \widetilde{\widetilde{Z}} \widetilde{z}_{ij} \widetilde{\widetilde{\xi}}_h n_{ijh}^{\widetilde{\alpha}-1}$$

  – For convenience when aggregating, we can express this in terms of shadow wages by multiplying through by the equilibrium multiplier on the rationing constraint

$$\begin{aligned} \underline{w} p_{ijh} &= p_{ijh} \widetilde{\alpha} \widetilde{\widetilde{Z}} \widetilde{z}_{ij} \widetilde{\widetilde{\xi}}_h n_{ijh}^{\widetilde{\alpha}-1} \\ \widetilde{w}_{ijh} &= p_{ijh} \times \widetilde{\alpha} \widetilde{\widetilde{Z}} \widetilde{z}_{ij} \widetilde{\widetilde{\xi}}_h n_{ijh}^{\widetilde{\alpha}-1} \end{aligned}$$

- **Region II - Constrained, on labor supply curve**

  – Now consider a constrained firm in Region II, this firm simply has labor determined by the labor supply curve, but since the rationing constraint is slack, $\widetilde{w}_{ijh} = p_{ijh} w_{ijh} = \underline{w}$.

$$n_{ijh} = \left(\frac{\underline{w}}{\widetilde{w}_{jh}} \right)^{\eta} \left(\frac{\widetilde{w}_{jh}}{\widetilde{W}_h} \right)^{\theta} n_h.$$

- Nonetheless, we can express the shadow wage of the firm as

$$\widetilde{w}_{ijh} = \widetilde{\mu}_{ijh}\widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}-1}$$
$$\widetilde{\mu}_{ijh} = \frac{w}{\widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}-1}} \quad , \quad n_{ijh} = \left(\frac{w}{\widetilde{w}_{jh}}\right)^{\eta}\left(\frac{\widetilde{w}_{jh}}{\widetilde{W}_h}\right)^{\theta} n_h.$$

- Therefore, in all three regions, we can express the *shadow wage* as a *shadow markdown* on the marginal revenue product of labor:

$$\widetilde{w}_{ijh} = \widetilde{\mu}_{ijh}\widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}-1}.$$

G.1.3 Aggregation of output and labor demand conditions

- Using the above results for firm optimality and the household's labor supply system we can aggregate the optimality conditions of agents. This is a key step in solving the government problem and optimal transfers, which we describe below.

- **Aggregation - Firm-Type to Market-Type**

  - From the above we have the following set of five conditions at the firm and market level:
  - **Firm level**:

$$\widetilde{y}_{ijh} = \widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}}$$
$$\widetilde{w}_{ijh} = \widetilde{\mu}_{ijh}\widetilde{\alpha}\widetilde{Z}\widetilde{z}_{ij}\widetilde{\xi}_h n_{ijh}^{\widetilde{\alpha}-1}.$$
$$n_{ijh} = \left(\frac{\widetilde{w}_{ijh}}{\widetilde{w}_{jh}}\right)^{\eta} n_{jh}$$

  - **Aggregates**:

$$\widetilde{y}_{jh} = \sum_{i\in j}\widetilde{y}_{ijh}$$
$$\widetilde{w}_{jh} = \left[\sum_{i\in j}\widetilde{w}_{ijh}^{1+\eta}\right]^{\frac{1}{1+\eta}}$$

  - Following steps from Berger, Herkenhoff, Mongey (2022), these can be combined to yield:

$$\begin{aligned}\widetilde{y}_{jh} &= \omega_{jh}\widetilde{Z}\widetilde{\xi}_h\widetilde{z}_j n_{jh}^{\widetilde{\alpha}}\\ \widetilde{w}_{jh} &= \widetilde{\mu}_{jh}\widetilde{\alpha}\widetilde{Z}\widetilde{z}_j\widetilde{\xi}_h n_{jh}^{\widetilde{\alpha}-1}\\ n_{jh} &= \left(\frac{\widetilde{w}_{jh}}{\widetilde{w}_h}\right)^{\theta} n_h\end{aligned}$$

where the three wedges $\{\widetilde{z}_j, \widetilde{\mu}_{jh}, \omega_{jh}\}$ are given by

$$\begin{aligned}\widetilde{z}_j &= \left[\sum_{i\in j}\widetilde{z}_{ij}^{\frac{1+\eta}{1+\eta(1-\widetilde{\alpha})}}\right]^{\frac{1+\eta(1-\widetilde{\alpha})}{1+\eta}}\\ \widetilde{\mu}_{jh} &= \left[\sum_{i\in j}\left(\frac{\widetilde{z}_{ij}}{\widetilde{z}_j}\right)^{\frac{1+\eta}{1+\eta(1-\widetilde{\alpha})}}\widetilde{\mu}_{ijh}^{\frac{1+\eta}{1+\eta(1-\widetilde{\alpha})}}\right]^{\frac{1+\eta(1-\widetilde{\alpha})}{1+\eta}}\\ \omega_{jh} &= \left[\sum_{i\in j}\left(\frac{\widetilde{z}_{ij}}{\widetilde{z}_j}\right)^{\frac{1+\eta}{1+\eta(1-\widetilde{\alpha})}}\left(\frac{\widetilde{\mu}_{ijh}}{\widetilde{\mu}_{jh}}\right)^{\frac{1+\eta}{1+\eta(1-\widetilde{\alpha})}}\right]^{\frac{1+\eta(1-\widetilde{\alpha})}{1+\eta}}\end{aligned}$$

- Note that this implies that if $\{\widetilde{z}_j, \widetilde{\mu}_{jh}, \widetilde{\omega}_{jh}\}$ are known, then $\{n_{jh}, \widetilde{w}_{jh}, \widetilde{y}_{jh}\}$ can be determined.

- **Aggregation - Market-Type to Type**
  - The same approach can be followed to aggregate to the household level, which delivers:

$$\begin{aligned}\widetilde{y}_h &= \omega_h\widetilde{Z}\widetilde{\xi}_h\widetilde{z}_h n_h^{\widetilde{\alpha}}\\ \widetilde{w}_h &= \widetilde{\mu}_h\widetilde{\alpha}\widetilde{Z}\widetilde{z}_h\widetilde{\xi}_h n_h^{\widetilde{\alpha}-1}\end{aligned}$$

where

$$\widetilde{z}_h = \left[\int \widetilde{z}_j^{\frac{1+\theta}{1+\theta(1-\widetilde{\alpha})}} dj \right]^{\frac{1+\theta(1-\widetilde{\alpha})}{1+\theta}}$$

$$\widetilde{\mu}_h = \left[\int \left(\frac{\widetilde{z}_{ij}}{\widetilde{z}_j} \right)^{\frac{1+\theta}{1+\theta(1-\widetilde{\alpha})}} \widetilde{\mu}_{jh}^{\frac{1+\theta}{1+\theta(1-\widetilde{\alpha})}} dj \right]^{\frac{1+\theta(1-\widetilde{\alpha})}{1+\theta}}$$

$$\omega_h = \left[\int \left(\frac{\widetilde{z}_{ij}}{\widetilde{z}_j} \right)^{\frac{1+\theta}{1+\theta(1-\widetilde{\alpha})}} \left(\frac{\widetilde{\mu}_{jh}}{\widetilde{\mu}_h} \right)^{\frac{1+\theta}{1+\theta(1-\widetilde{\alpha})}} \omega_{jh} \right]^{\frac{1+\theta(1-\widetilde{\alpha})}{1+\theta}}$$

- The conditions derived thus far all hold in a competitive equilibrium with lump sum transfers.
- In a competitive equilibrium, the above conditions are satisfied and budget constraints clear for each household.

I Mathematical details - Tax and Transfer System

In this appendix, we derive the household labor supply curve presented in Section 6.

Household. We adopt the Benabou (2003) and Heathcote et al (2017) tax and transfer system. We assume that after-tax income is a log-linear function of pre-tax income, where τ denotes the marginal tax rate and λ determines the level of transfers. We assume labor income per worker is w_{ij}, there are n_{ij} workers in the stand-in household, and thus after-tax income is $\left(\lambda w_{ij}^{1-\tau}\right) n_{ij}$. We present the economy with homogeneous workers, linear preferences over consumption, and omitting capital to minimize clutter. The household solves the following problem:

$$\max_{\{n_i\}_{i\in[0,1]}} U\left(C,N\right) = \frac{C^{1-\sigma}}{1-\sigma} - \frac{1}{\overline{\varphi}^{1/\varphi}}\frac{N^{1+1/\varphi}}{1+1/\varphi}$$

$$N = \left[\int_j n_j^{\frac{1+\theta}{\theta}} dj\right]^{\frac{\theta}{1+\theta}} \qquad n_j = \left[\sum_i n_{ij}^{\frac{\eta+1}{\eta}}\right]^{\frac{\eta}{\eta+1}}$$

subject to

$$C \leq \int \sum_i \left(\lambda w_{ij}^{1-\tau}\right) n_{ij} dj + \Pi \qquad , \qquad n_{ij} \leq \overline{n}_{ij}.$$

The corresponding Lagrangian is given by,

$$\mathcal{L} = \left[\frac{C^{1-\sigma}}{1-\sigma} - \frac{1}{\overline{\varphi}^{1/\varphi}}\frac{N^{1+1/\varphi}}{1+1/\varphi}\right] + \Lambda\left[\int_j \sum_i \left(\lambda w_{ij}^{1-\tau}\right) n_{ij} dj + \Pi - C\right] + \int_j \sum_i \phi_{ij}\left[\overline{n}_{ij} - n_{ij}\right] dj.$$

We then rewrite the *normalized* multiplier as follows:

$$\phi_{ij} = \left(1 - p_{ij}^{1-\tau}\right)\Lambda\lambda w_{ij}^{1-\tau}.$$

After substituting for ϕ_{ij}, collecting terms, and letting let $\widetilde{w}_{ij} := w_{ij}p_{ij}$, we have the following Lagrangian:

$$\mathcal{L} = \left[\frac{C^{1-\sigma}}{1-\sigma} - \frac{1}{\overline{\varphi}^{1/\varphi}}\frac{N^{1+1/\varphi}}{1+1/\varphi}\right] + \Lambda\left[\int_j \sum_i \lambda \widetilde{w}_{ij}^{1-\tau} n_{ij} dj + \Pi - C\right] + \int_j \sum_i \left(1 - p_{ij}^{1-\tau}\right)\Lambda\lambda w_{ij}^{1-\tau}\overline{n}_{ij} dj.$$

The FOC for consumption yields

$$U_C = \Lambda.$$

The FOC for n_{ij} yields

$$-U_N \frac{\partial N}{\partial n_j} \frac{\partial n_j}{\partial n_{ij}} = U_C \lambda \widetilde{w}_{ij}^{1-\tau}$$

We define $\widetilde{w}_j$ and $\widetilde{W}$ such that

$$\lambda \widetilde{w}_j^{1-\tau} n_j = \sum_{i \in j} \lambda \widetilde{w}_{ij}^{1-\tau} n_{ij} \qquad , \qquad \lambda \widetilde{W}^{1-\tau} N = \sum_j \lambda \widetilde{w}_j^{1-\tau} n_j$$

Aggregating to the household-level the first order condition for n_{ij},we can then solve for the labor supply curve of representative household

$$-\frac{U_N N}{U_C} = \lambda \widetilde{W}^{1-\tau} N.$$

Then substituting this back into our first order condition for n_{ij} and aggregating to the market-level yields:

$$n_j = \left(\frac{\widetilde{w}_j}{\widetilde{W}} \right)^{\theta(1-\tau)} N$$

Substituting back into n_{ij}yields the inverse labor supply curve:

$$\widetilde{w}_{ij} = \left(\frac{n_{ij}}{n_j} \right)^{\frac{1}{(1-\tau)\eta}} \left(\frac{n_j}{N} \right)^{\frac{1}{(1-\tau)\theta}} \widetilde{W}$$

Inversion yields the labor supply curve shown in the text:

$$n_{ij} = \left(\frac{\widetilde{w}_{ij}}{\widetilde{w}_j} \right)^{(1-\tau)\eta} \left(\frac{\widetilde{w}_j}{\widetilde{W}} \right)^{\theta(1-\tau)} N$$

Firm. Abstracting from capital, the firm problem is to choose n_{ij},and the rationing constraint $\overline{n}_{ij}$ to maximize profits

$$\pi_{ij} = \max_{n_{ij}, \overline{n}_{ij}} z_{ij} n_{ij}^{\alpha} - w\left(n_{ij}\right) n_{ij},$$

where $w\left(n_{ij}\right)$ is the household labor supply curve under HSV taxes, $w\left(n_{ij}\right) \geq \underline{w}$, and $n_{ij} \leq \overline{n}_{ij}$.

Characterization. Dividing the firm problem into three regions, we arrive at a similar characterization of wage setting as our baseline economy.

Region I wages and allocations:

$$w_{ij} = \frac{\varepsilon_{ij}}{\varepsilon_{ij}+1} \alpha z_{ij} n_{ij}^{\alpha-1}, \qquad \varepsilon_{ij}^{-1} = \frac{1}{1-\tau}\left[\frac{1}{\eta} + \left(\frac{1}{\theta} - \frac{1}{\eta}\right)\left(\frac{w_{ij}}{\widetilde{w}_j}\right)^{(1+\eta)(1-\tau)}\right],$$
$$n_{ij} = \left(\frac{w_{ij}}{\widetilde{w}_j}\right)^{(1-\tau)\eta} \left(\frac{\widetilde{w}_j}{\widetilde{W}}\right)^{\theta(1-\tau)} N, \qquad p_{ij} = 1$$

Region II wages and allocations:

$$w_{ij} = \underline{w}, \qquad n_{ij} = \left(\frac{\underline{w}}{\widetilde{w}_j}\right)^{(1-\tau)\eta} \left(\frac{\widetilde{w}_j}{\widetilde{W}}\right)^{\theta(1-\tau)} N, \qquad p_{ij} = 1$$

Region III wages and allocations:

$$w_{ij} = \underline{w}, \qquad n_{ij} = \left(\frac{\alpha z_{ij}}{\underline{w}}\right)^{1/(1-\alpha)}, \qquad p_{ij} = \frac{\left(\frac{n_{ij}}{n_j}\right)^{\frac{1}{(1-\tau)\eta}} \left(\frac{n_j}{N}\right)^{\frac{1}{(1-\tau)\theta}} \widetilde{W}}{\underline{w}}$$

where the wage indeces are given by,

$$\widetilde{w}_{ij} = p_{ij} w_{ij}, \qquad \widetilde{w}_j = \left[\sum_{i \in j} \widetilde{w}_{ij}^{(1+\eta)(1-\tau)}\right]^{\frac{1}{(1+\eta)(1-\tau)}}, \qquad \widetilde{W} = \left[\sum_j \widetilde{w}_j^{(1+\eta)(1-\tau)}\right]^{\frac{1}{(1+\eta)(1-\tau)}}.$$

The solution algorithm proceeds in an identical manner to our baseline economy.

I Optimal Tax and Transfer System

We thank an anonymous referee for suggesting this exercise. We compute the total potential welfare gains from changes in the tax and transfer policy in an economy with $\underline{w} = 0$. We limit our exercise to optimizing the Heathcote, Storesletten, and Violante (2017) (henceforth, HSV) tax function. Mechanically, we proceeds as follows:

1. The tax policy consists of three parameters: τ progressivity, λ the shifter on the tax function which determines the point at which it goes from a subsidy to a tax, and g the implied share of output that is accounted for by the net government spending position of the tax and transfer system.

2. We choose (τ, λ) to match the data and recalibrated the parameters of the economy

Figure I1: Optimal Progressivity

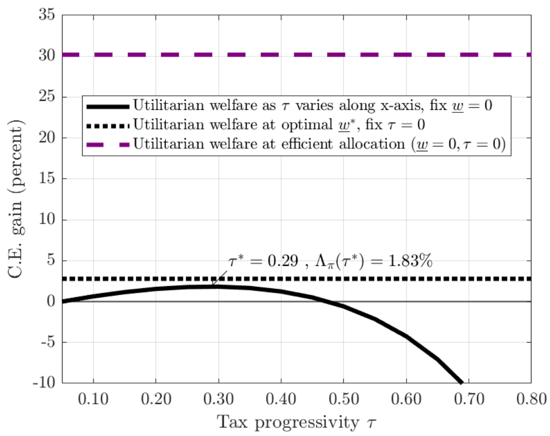
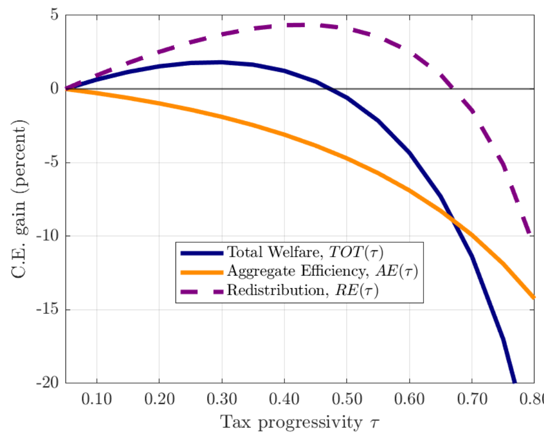


Figure I2: Efficiency and Redistribution



so that we match the same wage moments *pre-tax*, and backed out the implied g, which was small and positive.

3. We then *adjust the progressivity of the tax system*. We consider alternative values of $\tau' \in [0.05, 0.80]$, computing for each τ' the required change in λ' to deliver the same balance g. In this case, raising τ raises more taxes on high income workers, which allows for an expansion of the cut-off of the policy and expansion of the maximum transfer.

Results. Figure I1 summarizes our main result. The optimal degree of progressivity and subsidy/tax cut-off are $\tau^* = 0.29$ and $\lambda^* = 2.39$. This is more progressivity than the empirical baseline of $\tau = 0.18$, and a larger threshold for receipt of a net subsidy than the empirical baseline of $\lambda = 1.74$. The overall welfare gain relative to the empirical (τ, λ) is 1.83% to the Utilitarian planner. This remains small relative to the gains an unrestricted planner would achieve of 30.2%. Relative to the baseline (τ, λ) we find that the gains from (τ^*, λ^*) decompose into roughly a 4% *Redistribution* gain and 3% *Efficiency* loss. See Figure I2.

Discussion. Why are the welfare gains from optimizing the tax policy parameters—subject to the fiscal position of the government—small relative to the welfare gains a planner could achieve? There are two main reasons. First, the empirical distribution of consumption, labor, labor income and non-labor income in the economy—which the model is calibrated to— is far from what a Utilitarian planner would choose. For example, 33 per-

Figure I3: C_h for $\tau = 0.18$ vs $\tau^* = 0.29$

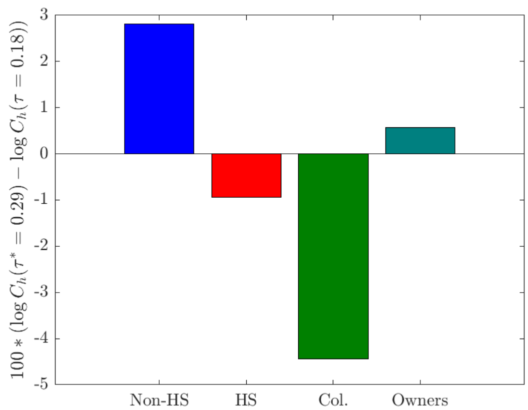
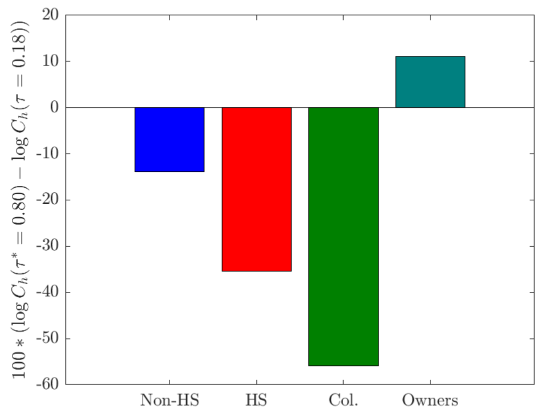


Figure I4: C_h for $\tau = 0.18$ vs $\tau' = 0.80$



cent of households have a college degree, yet they account for 72 percent of consumption. A tax and transfer system that maximizes a Utilitarian objective would involve massive transfers away from these households. Unrestricted, subsidies are going to look nothing like tax policy that we see in the data.

Second, a higher τ widens markdowns as firms internalize it being more expensive to hire labor on the margin. This can then generate efficiency losses. For $\tau = 0.29$, there is redistribution toward lower income households which would be the 'intended' effect of the policy (see Figure I3). As τ increases further, the 'unintended consequences' kick in limiting the redistributive gains. In Figure I4, we consider $\tau = 0.80$. At this level markdowns are wider and profits are flowing to the highest income business owners. Hence on top of the lower efficiency, redistributive gains are now undone.

Caveats. First, we don't know what the correct social welfare weights are and results depend critically on welfare weights.[52] Second, and related, the U.S. equilibrium—which we calibrate our model to—is inconsistent with an allocation of resources that a social planner with Utilitarian weights would choose. To maximize the *redistribution component* of welfare, requires massively increasing the consumption of the very large mass of high-school and non-high-school educated households in the economy that consume very little. An EITC/HSV tax function is not going to achieve this, only a massive overhaul of the tax and transfer system. Third, the tax and transfer system is not just about redistribution, but

[52] That the competitive equilibrium is so far from the Utilitarian optimal should give us some pause that Utilitarian weights are not the right ones for policy analysis, but we don't know what the correct ones are either.

Figure J1: Lower elasticities $(\eta', \theta') = (0.7 \times \eta, 0.7 \times \theta)$

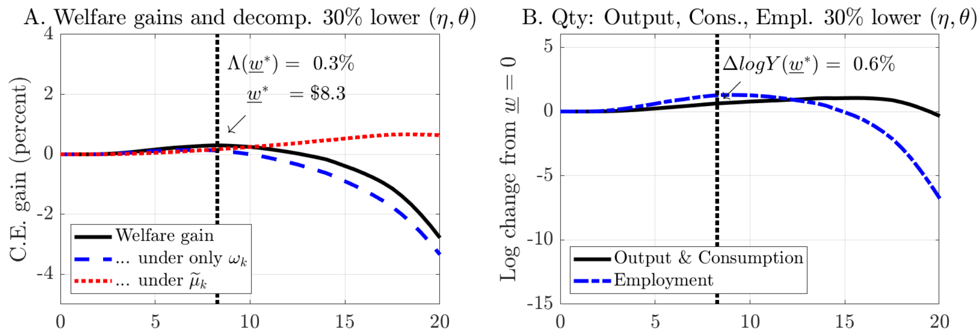
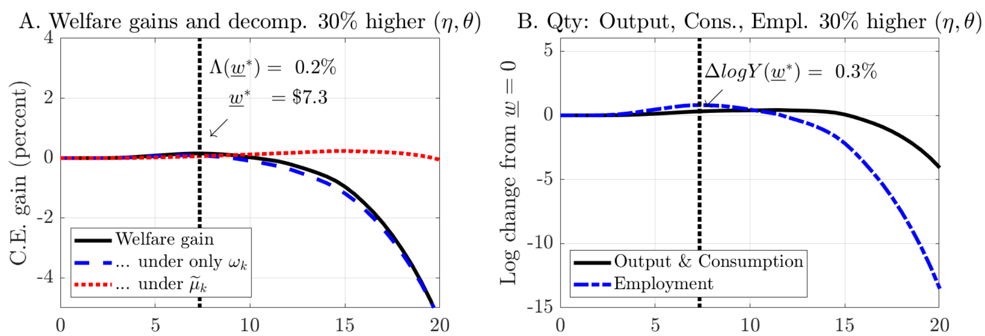


Figure J2: Higher elasticities $(\eta', \theta') = (1.3 \times \eta, 1.3 \times \theta)$



also insurance. Heathcote, Storesletten, and Violante (2017) show persuasively that in an economy with idiosyncratic risk, the optimal degree of progressivity—in a formulation of taxes that we use in this paper—increases welfare primarily through providing insurance. Our economy does not have such idiosyncratic risk, and hence abstracts from this key benefit.

J Robustness to η and θ

Higher (Figure I6) and lower (Figure I5) values of labor supply elasticities do very little to the magnitude of efficiency gains. Decreasing (Increasing) elasticities by 30% increases (decreases) the optimal minimum wage by 70c (30c), and leaves welfare gains almost unchanged. The intuition is that firm productivity dispersion limits the effectiveness of minimum wages, even when markdowns are greater at smaller firms (see Figure I5).