Online Appendix

Markups and Inequality

Not For Publication

Corina Boar[1] Virgiliu Midrigan[2]

November 2023

This appendix provides detailed derivations for our optimal regulation results in Section 2, reports results from an economy in which firms have labor market power, describes the SCF data we used to parameterize the model, and reports the details of the robustness exercises reported in the main text for both the static and dynamic model.

1 Optimal Regulation in the Baseline Model

We first derive the incentive compatibility constraint and then solve the regulator's problem.

1.1 Incentive Compatibility Constraints

We assume that the regulator does not observe entrepreneurial ability z. Without loss of generality, we invoke the revelation principle and constrain the regulator to choose functions $c(z)$ and $q(z)$ that ensure truth-telling.

Let $\tau(z)$ denote a subsidy received by a firm that claims to be of type z and sells $q(z) \equiv y(z)/Y$ units of output as prescribed by the regulator. The producer's consumption, if it reports truth-fully, is

$$c(z) = DY\left[\Upsilon'(q(z))\,q(z) - \frac{W(z)\,Y^{\frac{1}{\eta}-1}}{D}\left(\frac{q(z)}{z}\right)^{\frac{1}{\eta}} + \tau(z)\right].$$

If the producer claims instead to have productivity $\hat{z}$, it receives

$$\hat{c}(\hat{z}, z) = DY\left[\Upsilon'(q(\hat{z}))\,q(\hat{z}) - \frac{W(z)\,Y^{\frac{1}{\eta}-1}}{D}\left(\frac{q(\hat{z})}{z}\right)^{\frac{1}{\eta}} + \tau(\hat{z})\right]$$

[1]New York University, and NBER, corina.boar@nyu.edu.
[2]New York University, and NBER, virgiliu.midrigan@nyu.edu.

units of consumption. Following the first-order approach, the local incentive constraint is

$$\left.\frac{\partial \hat{c}\left(\hat{z}, z\right)}{\partial \hat{z}}\right|_{\hat{z}=z} = 0 = DY\left[\left(\Upsilon''\left(q\left(z\right)\right)q\left(z\right) + \Upsilon'\left(q\left(z\right)\right) - \frac{1}{\eta}\frac{W\left(z\right)Y^{\frac{1}{\eta}-1}}{D}\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}\frac{1}{q\left(z\right)}\right)q'\left(z\right) + \tau'\left(z\right)\right].$$

Differentiating the expression for $c(z)$ with respect to z and imposing the local incentive constraint gives

$$c'\left(z\right) = Y^{\frac{1}{\eta}}\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}\frac{W\left(z\right)}{z}\left[\frac{1}{\eta} - \frac{W'\left(z\right)z}{W\left(z\right)}\right].$$

1.2 Regulator's Problem

In Section 2.4 we formulated the regulator's problem as choosing $y(z)$, $c(z)$, Y, W_1 and W_2. We can equivalently recast the regulator's problem as choosing relative output $q(z) = y(z)/Y$, $c(z)$, Y, W_1 and the skill premium $S = W_2/W_1$. Let

$$\mathcal{W}\left(z, S\right) = \frac{W\left(z\right)}{W_1} = \left[\left(1-\psi\left(z\right)\right) + \psi\left(z\right)S^{1-\rho}\right]^{\frac{1}{1-\rho}}$$

denote the composite wage of a firm with productivity z relative to the wage of low-skill workers, W_1. This object only depends on the skill premium S and firm productivity. With this notation, the labor market clearing conditions can then be written as

$$L_1^w(W_1) = \omega Y^{\frac{1}{\eta}}\int\left(1-\psi\left(z\right)\right)\mathcal{W}\left(z, S\right)^{\rho}\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}\mathrm{d}F\left(z\right)$$

and

$$L_2^w(W_2) = \omega S^{-\rho}Y^{\frac{1}{\eta}}\int\psi\left(z\right)\mathcal{W}\left(z, S\right)^{\rho}\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}\mathrm{d}F\left(z\right),$$

and the incentive compatibility constraint is

$$c'\left(z\right) = Y^{\frac{1}{\eta}}\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}\frac{W_1}{z}\left[\frac{\mathcal{W}\left(z, S\right)}{\eta} - \frac{\zeta}{1-\rho}\left(S^{1-\rho}-1\right)\mathcal{W}\left(z, S\right)^{\rho}\left(1-\psi\left(z\right)\right)\psi\left(z\right)\right].$$

The Lagrangean is

$$\max_{q(z),\,c(z),\,Y,\,W_1,\,S} V_1^w\left(W_1\right)+V_2^w\left(W_1 S\right)+\omega\int\frac{c\left(z\right)^{1-\theta}}{1-\theta}f\left(z\right)\mathrm{d}z$$
$$+\int\hat{\mu}\left(z\right)\left[c'\left(z\right)-Y^{\frac{1}{\eta}}\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}\frac{W_1}{z}\left(\frac{\mathcal{W}\left(z,S\right)}{\eta}-\frac{\zeta}{1-\rho}\left(S^{1-\rho}-1\right)\mathcal{W}\left(z,S\right)^{\rho}\left(1-\psi\left(z\right)\right)\psi\left(z\right)\right)\right]\mathrm{d}z$$
$$+\lambda\left[Y-C_1^w\left(W_1\right)-C_2^w\left(W_1 S\right)-\omega\int c\left(z\right)f\left(z\right)\mathrm{d}z\right]+\kappa\left[\omega\int\Upsilon\left(q\left(z\right)\right)f\left(z\right)\mathrm{d}z-1\right]$$
$$+\hat{\nu}_1\left[L_1\left(W_1\right)-Y^{\frac{1}{\eta}}\omega\int\left(1-\psi\left(z\right)\right)\mathcal{W}\left(z,S\right)^{\rho}\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}f\left(z\right)\mathrm{d}z\right]$$
$$+\hat{\nu}_2\left[L_2\left(W_1 S\right)-S^{-\rho}Y^{\frac{1}{\eta}}\omega\int\psi\left(z\right)\mathcal{W}\left(z,S\right)^{\rho}\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}f\left(z\right)\mathrm{d}z\right],$$

where $\hat{\mu}\left(z\right)$ are the multipliers on the IC constraints and $\hat{\nu}_s$ denote the multipliers on the labor resource constraints.

Consider the term $\int_0^\infty\hat{\mu}\left(z\right)c'\left(z\right)\mathrm{d}z$. Integrating by parts and using the boundary conditions $\hat{\mu}\left(0\right)=\hat{\mu}\left(\infty\right)=0$ gives[3]

$$\int_0^\infty\hat{\mu}\left(z\right)c'\left(z\right)\mathrm{d}z=-\int_0^\infty\hat{\mu}'\left(z\right)c\left(z\right)\mathrm{d}z,$$

which, after rearranging, allows us to rewrite the Lagrangean as

$$\max_{q(z),\,c(z),\,Y,\,W_1,\,S} V_1^w\left(W_1\right)+V_2^w\left(W_1 S\right)+\omega\int\frac{c\left(z\right)^{1-\theta}}{1-\theta}f\left(z\right)\mathrm{d}z-\int\hat{\mu}'\left(z\right)c\left(z\right)\mathrm{d}z$$

$$-Y^{\frac{1}{\eta}}\omega\int\underbrace{W_1\frac{\mu\left(z\right)}{\omega f\left(z\right)}\frac{1}{z}\left[\frac{\mathcal{W}\left(z,S\right)}{\eta}-\frac{\zeta}{1-\rho}\left(S^{1-\rho}-1\right)\mathcal{W}\left(z,S\right)^{\rho}\left(1-\psi\left(z\right)\right)\psi\left(z\right)\right]}_{\hat{B}_1(z)}\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}f\left(z\right)\mathrm{d}z$$

$$+\lambda\left[Y-C_1^w\left(W_1\right)-C_2^w\left(W_1 S\right)-\omega\int c\left(z\right)f\left(z\right)\mathrm{d}z\right]+\kappa\left[\omega\int\Upsilon\left(q\left(z\right)\right)f\left(z\right)\mathrm{d}z-1\right]$$

$$+\hat{\nu}_1 L_1\left(W_1\right)+\hat{\nu}_2 L_2\left(W_1 S\right)-Y^{\frac{1}{\eta}}\omega\int\underbrace{\left(\hat{\nu}_1\left(1-\psi\left(z\right)\right)+\hat{\nu}_2 S^{-\rho}\psi\left(z\right)\right)\mathcal{W}\left(z,S\right)^{\rho}}_{\hat{B}_2(z)}\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}f\left(z\right)\mathrm{d}z.$$

[3] There are no distortions at the top because in our numerical experiments we bound the distribution of ability.

Notice that the terms $\hat{B}_1(z)$ and $\hat{B}_2(z)$ can be more compactly written as

$$\hat{B}_1(z) = \frac{\hat{\mu}(z)}{\omega f(z)} \frac{W(z)}{z} \left[\frac{1}{\eta} - \frac{W'(z) z}{W(z)} \right]$$

and

$$\hat{B}_2(z) = \left(\hat{\nu}_1 W_1^{-\rho} (1 - \psi(z)) + \hat{\nu}_2 W_2^{-\rho} \psi(z) \right) W(z)^{\rho},$$

a formulation that we will use below.

To solve the regulator's problem, we now list the first-order conditions with respect to each of the choice variables. The FOC with respect to $c(z)$ implies

$$\hat{\mu}'(z) = \left(\alpha c(z)^{-\theta} - \lambda \right) \omega f(z).$$

The FOC with respect to $q(z)$ is

$$\kappa \Upsilon'(q(z)) \omega f(z) = \frac{1}{\eta} Y^{\frac{1}{\eta}} \omega f(z) \left[\hat{B}_1(z) + \hat{B}_2(z) \right] \left(\frac{q(z)}{z} \right)^{\frac{1}{\eta}} \frac{1}{q(z)}.$$

The FOC with respect to Y is

$$\lambda = \frac{1}{\eta} Y^{\frac{1}{\eta}-1} \omega \int \left[\hat{B}_1(z) + \hat{B}_2(z) \right] \left(\frac{q(z)}{z} \right)^{\frac{1}{\eta}} f(z)\, \mathrm{d}z = \frac{\kappa}{DY},$$

where the last equality follows from integrating the $q(z)$ FOC and using the definition of D. The FOC with respect to W_1 is

$$\begin{aligned} V_1'(W_1) + V_2'(W_2) S - \lambda \left[C_1'(W_1) + C_2'(W_2) S \right] + \hat{\nu}_1 L_1'(W_1) + \hat{\nu}_2 L_2(W_2) S \\ = \frac{1}{W_1} Y^{\frac{1}{\eta}} \omega \int \hat{B}_1(z) \left(\frac{q(z)}{z} \right)^{\frac{1}{\eta}} f(z)\, \mathrm{d}z. \end{aligned}$$

Lastly, the FOC with respect to S is

$$V_2'(W_2) W_1 - \lambda C_2'(W_2) W_1 + \hat{\nu}_2 L_2'(W_2) W_1 = Y^{\frac{1}{\eta}} \omega \int \left[\frac{\partial \hat{B}_1(z,S)}{\partial S} + \frac{\partial \hat{B}_2(z,S)}{\partial S} \right] \left(\frac{q(z)}{z} \right)^{\frac{1}{\eta}} f(z)\, \mathrm{d}z.$$

To derive equation (21) in the text that determines the regulator's optimal choice of output across producers, we note that, since $\hat{\mu}(\infty) = 0$,

$$\hat{\mu}(z) = - \int_z^{\infty} \hat{\mu}'(x)\, dx = \omega \int_z^{\infty} \left[\lambda - c(x)^{-\theta} \right] f(x)\, \mathrm{d}x.$$

Moreover, since $\hat{\mu}(0) = 0$ we have

$$\hat{\mu}(0) = \omega \int_0^\infty \left[\lambda - c(z)^{-\theta}\right] f(z)\,\mathrm{d}z = 0,$$

which implies that

$$\lambda = \int_0^\infty c(z)^{-\theta} f(z)\,\mathrm{d}z.$$

The $q(z)$ and Y FOCs together imply

$$\Upsilon'(q(z))\,q(z) = \left(\hat{B}_1(z) + \hat{B}_2(z)\right) \frac{1}{\eta} \frac{Y^{\frac{1}{\eta}-1}}{\lambda D} \left(\frac{q(z)}{z}\right)^{\frac{1}{\eta}} \tag{1}$$

Let

$$\mu(z) \equiv \frac{\hat{\mu}(z)}{\omega\lambda(1 - F(z))},$$

and

$$\nu_s \equiv \frac{\hat{\nu}_s}{\lambda},$$

as well as

$$B_1(z) \equiv \frac{\hat{B}_1(z)}{\lambda W(z)} = \frac{\mu(z)(1 - F(z))}{f(z)} \frac{1}{z} \left[\frac{1}{\eta} - \frac{W'(z)\,z}{W(z)}\right]$$

and

$$B_2(z) \equiv \frac{\hat{B}_2(z)}{\lambda W(z)} = \frac{(1 - \psi(z))\,\nu_1 W_1^{-\rho} + \psi(z)\,\nu_2 W_2^{-\rho}}{(1 - \psi(z))\,W_1^{1-\rho} + \psi(z)\,W_2^{1-\rho}.}$$

We can now rewrite equation (1) in terms of output $y(z)$ as follows

$$\begin{aligned} D\Upsilon'\left(\frac{y(z)}{Y}\right) &= (B_1(z) + B_2(z)) \frac{W(z)}{\eta} \left(\frac{y(z)}{z}\right)^{\frac{1}{\eta}} \frac{1}{y(z)} \\ &= \left(1 + \frac{B_1(z)}{B_2(z)}\right) B_2(z) \frac{W(z)}{\eta} \left(\frac{y(z)}{z}\right)^{\frac{1}{\eta}} \frac{1}{y(z)} \end{aligned} \tag{2}$$

Letting

$$\nu(z) \equiv \frac{\hat{B}_2(z)}{\lambda}$$

denote the effective marginal rate of substitution between composite employment at a firm with productivity z and consumption, we have

$$B_2(z) = \frac{\nu(z)}{W(z)}$$

and

$$\frac{B_1(z)}{B_2(z)} = \frac{\mu(z)(1-F(z))}{\nu(z) z f(z)} \left[\frac{W(z)}{\eta} - W'(z) z\right].$$

Using this and rearranging equation (2) gives the equation (21) in the text.

Given quantities $q(z)$, the ICC implies that the consumption of each entrepreneur,

$$c(z) = c(0) + W_1 Y^{\frac{1}{\eta}} \int_0^z \left(\frac{q(x)}{x}\right)^{\frac{1}{\eta}} \frac{1}{x} \left[\frac{\mathcal{W}(x)}{\eta} - \frac{\zeta}{1-\rho}\left(S^{1-\rho}-1\right) \mathcal{W}(x)^{\rho} (1-\psi(x)) \psi(x)\right] \mathrm{d}x,$$

where the lump-sum transfer $c(0)$ adjusts to ensure revenue neutrality

$$C_1^w(W_1) + C_2^w(W_2) + \omega \int c(z) f(z) \,\mathrm{d}z = Y.$$

We solve this problem iteratively, by first conjecturing the optimal quantities $q(z)$, and then solving the system of equations that, together with the incentive compatibility constraints, determine λ, ν_1, ν_2, W_1, S, Y and D. Given these equilibrium objects, we obtain the wedges $\xi(z)$ and $\nu(z)/W(z)$ and update the guess for $q(z)$ using equation (21) in the text until convergence. After the algorithm converges we verify numerically that the global incentive constraints are satisfied.

2 An Economy with Both Goods and Labor Market Power

Here we describe an economy in which firms not only have market power in the product market, as in our baseline economy, but also monopsony power in the labor market. For expositional clarity, we abstract from skill heterogeneity. We present the setup of the model, solve the regulator's problem and report results from the optimal policy experiments. We show that our conclusions that optimal regulation leads to higher product market concentration are robust to the sources of firm market power.

2.1 Environment

The economy is inhabited by a measure ω of entrepreneurs and a measure $1-\omega$ of workers.

2.1.1 Workers

Workers have preferences

$$u(c,h) = \frac{c^{1-\theta}}{1-\theta} - \frac{h^{1+\gamma}}{1+\gamma},$$

where h now denotes the composite amount of labor supplied to all firms and depends on the amount of hours h_i supplied to each firm i according to a Kimball aggregator

$$\int_0^\omega \Upsilon_h \left(\frac{h_i}{h}\right) \mathrm{d}i = 1.$$

As in our baseline model, consumers differ in their ability e so their budget constraint is

$$c = Weh,$$

where W is the aggregate wage index, implicitly defined by

$$Wh = \int_0^\omega w_i h_i \mathrm{d}i,$$

and w_i is the wage paid by firm i. Solving the worker's problem gives the amount of labor it supplies to each firm

$$\frac{w_i}{W} = D_h \Upsilon_h' \left(\frac{h_i}{h}\right), \tag{3}$$

where

$$D_h \equiv \left(\int_0^\omega \Upsilon_h' \left(\frac{h_i}{h}\right) \frac{h_i}{h} \mathrm{d}i\right)^{-1}.$$

Because preferences are homothetic, equation (3) also describes the aggregate supply of labor that firm i faces. That is, letting

$$l_i = \int h_i\left(e\right) dH\left(e\right)$$

and

$$L = \int h\left(e\right) dH\left(e\right),$$

each firm i faces the following upward-sloping labor supply equation

$$\frac{w_i}{W} = D_h \Upsilon_h' \left(\frac{l_i}{L}\right).$$

We note that the setting we assume here is closely related to that studied in Berger et al. (2022), except that we assume monopsonistic competition with a continuum of firms, rather than oligopsonistic competition with a finite number of firms, and a Kimball-style aggregator of the disutility from work. As we discuss below, the non-constant elasticity aggregator allows us to capture the idea that larger firms face a less elastic labor supply and thus enjoy more labor market power.

As in the baseline model, the equilibrium hours worked are

$$h(e,W) = (We)^{\frac{1-\theta}{\gamma+\theta}} \quad \text{and} \quad c(e,W) = (We)^{\frac{1+\gamma}{\gamma+\theta}}$$

and the welfare of workers is increasing in the aggregate wage index W

$$v(e,W) = u(c(e,W), h(e,W)) = \frac{\gamma+\theta}{(1-\theta)(1+\gamma)} W^{\frac{(1-\theta)(1+\gamma)}{\gamma+\theta}} e^{\frac{(1-\theta)(1+\gamma)}{\gamma+\theta}}.$$

2.1.2 Entrepreneurs

Entrepreneurs have preferences

$$u(c) = \frac{c^{1-\theta}}{1-\theta}$$

and operate the technology

$$y = zl^{\eta}$$

Their budget constraint is

$$c = \pi = p(y)y - w(l)l,$$

where $p(y)$ is the inverse demand function described in the main text and $w(l)$ is the inverse labor supply function described above. In contrast to the baseline model, the firm now recognizes that hiring more workers requires paying them a higher wage.

The profit maximization problem of entrepreneur i is

$$\max_{y_i, l_i} p_i y_i - w_i l_i = D\Upsilon'\left(\frac{y_i}{y}\right) y_i - WD_h \Upsilon_h'\left(\frac{l_i}{l}\right) l_i$$

or substituting the production function

$$\max_{y_i} D\Upsilon'\left(\frac{y_i}{y}\right) y_i - WD_h \Upsilon_h'\left(\frac{(y_i/z_i)^{\frac{1}{\eta}}}{l}\right)\left(\frac{y_i}{z_i}\right)^{\frac{1}{\eta}}.$$

The y_i FOC is

$$D\Upsilon''\left(\frac{y_i}{y}\right)\frac{y_i}{y} + D\Upsilon'\left(\frac{y_i}{y}\right) = \frac{1}{\eta}\left(\frac{y_i}{z_i}\right)^{\frac{1}{\eta}}\frac{1}{y_i}\left(WD_h\Upsilon_h''\left(\frac{(y_i/z_i)^{\frac{1}{\eta}}}{l}\right)\frac{(y_i/z_i)^{\frac{1}{\eta}}}{l} + WD_h\Upsilon_h'\left(\frac{(y_i/z_i)^{\frac{1}{\eta}}}{l}\right)\right).$$

Let $q_i = y_i/Y$ denote the firm's relative output and $x_i = l_i/L$ denote its relative employment. Then we can rewrite the FOC above as

$$\left[1 + \frac{\Upsilon''(q_i)\,q_i}{\Upsilon'(q_i)}\right] p_i = \frac{1}{\eta}\left(\frac{y_i}{z_i}\right)^{\frac{1}{\eta}} \frac{w_i}{y_i}\left[1 + \frac{\Upsilon_h''(x_i)\,x_i}{\Upsilon_h'(x_i)}\right]$$

Letting

$$m\left(q_i\right)=\left(1+\frac{\Upsilon''\left(q_i\right)q_i}{\Upsilon'\left(q_i\right)}\right)^{-1}\geq 1$$

and

$$m_h\left(x_i\right)=\left(1+\frac{\Upsilon_h''\left(x_i\right)x_i}{\Upsilon_h'\left(x_i\right)}\right)^{-1}\leq 1,$$

we can write the firm's price as a markup $m\left(q_i\right)$ (ratio of price to marginal cost) divided by a markdown $m_h\left(x_i\right)$ (ratio of wage to marginal revenue product of labor) times the firm's marginal cost.

$$p_i=\frac{m\left(q_i\right)}{m_h\left(x_i\right)}\times\frac{1}{\eta}\left(\frac{y_i}{z_i}\right)^{\frac{1}{\eta}}\frac{w_i}{y_i}$$

To see the model's implications for the labor share, note that the labor share is given by

$$\frac{w_i l_i}{p_i y_i}=\eta\frac{m_h\left(x_i\right)}{m\left(q_i\right)}.$$

Absent market power in the labor and the goods market, the labor share would be equal to η. Both markups and markdowns reduce the labor share. Moreover, if markups increase with firm size and markdowns fall with firm size, larger firms have a lower labor share. We also note that the production approach to estimating markups cannot separately distinguish markups from markdowns with only data on the labor share because these two wedges are observationally equivalent: larger firms may have a lower labor share either because they charge higher markups or because they pay their workers a lower markdown or both. As we show below, however, the two types of distortions lead to similar conclusions regarding optimal product market interventions.

To capture the possibility that markdowns decrease with firm relative employment, we assume a Kimball aggregator for the disutility from work of the form

$$\Upsilon_h(x)=1+(\nu+1)\exp\left(-\frac{1}{\epsilon}\right)\epsilon^{\frac{\nu}{\epsilon}-1}\left(-1\right)^{-\frac{\nu}{\epsilon}}\left[\Gamma\left(\frac{\nu}{\epsilon},-\frac{1}{\epsilon}\right)-\Gamma\left(\frac{\nu}{\epsilon},-\frac{x^{\epsilon/\nu}}{\epsilon}\right)\right],$$

which implies that the labor supply elasticity

$$\nu\left(x\right)\equiv\left(\frac{\Upsilon_h''\left(x\right)x}{\Upsilon_h'\left(x\right)}\right)^{-1}=\nu x^{-\frac{\epsilon}{\nu}}$$

decreases with the firm's relative employment x at a rate that depends on the super-elasticity ϵ/ν.

2.2 Optimal Regulation

To isolate the role of markdowns in shaping optimal regulation, we study separately two economies: one with only markups ($\sigma = \infty$) and another with only markdowns ($\nu = \infty$). Using the same approach we used to solve the utilitarian regulator's problem in the baseline economy, we can show that the optimal output allocation across producers in the economy with only markups is given by

$$\Upsilon'\left(q\left(z\right)\right)q\left(z\right) = \left(1 + \mu\left(z\right)\frac{\frac{1}{\eta}\frac{W}{z}\left(1 - F\left(z\right)\right)}{\nu f\left(z\right)}\right)\left(\frac{q\left(z\right)}{z}\right)^{\frac{1}{\eta}}\Omega, \tag{4}$$

where $\mu(z)$ is the same as in the baseline model and Ω is only a function of aggregate variables and is pinned down by the requirement that $\omega\int\Upsilon(q(z))\mathrm{d}F(z) = 1$. Note that this expression is the one we derived in Boar and Midrigan (2022).

Similarly, the optimal quantity prescribed by the utilitarian regulator in an economy with only markdowns is given by

$$\Upsilon_h'\left(x\left(z\right)\right)x\left(z\right) = \left[1 - \frac{\mu\left(z\right)}{f\left(z\right)}\frac{\left(1 - F\left(z\right)\right)}{z}\right]zx\left(z\right)^{\eta}\Omega_h,$$

where Ω_h is only a function of aggregate variables and is pinned down by the requirement that $\omega\int\Upsilon_h(x(z))\mathrm{d}F(z) = 1$. Relative to the efficient allocations, which satisfy

$$\Upsilon_h'\left(x\left(z\right)\right)x\left(z\right) = zx\left(z\right)^{\eta}\Omega_h,$$

the regulator introduces a wedge that, as in the economy with markups, captures the equity-efficiency tradeoff.

2.3 Results

We next discuss how we parameterize the two economies and report results from the optimal policy experiments.

2.3.1 Parameterization

We calibrate the economy with markups using the same strategy as in the baseline economy. The only difference is that, since we only have one type of workers, we no longer target the skill premium and the wage-employment elasticity. Table 1 reports the parameter values and Table 2 reports the fit of the model.

We calibrate the economy with markdowns using a similar strategy, with one exception: we also calibrate the super-elasticity of labor supply ϵ/ν to reproduce an elasticity of firm

Table 1: Parameter Values in Economy with Markups

Assigned			Calibrated		
θ	1	CRRA coefficient	σ_e	1.00	std. dev. Gaussian term, workers
γ	2	inverse Frisch elasticity	λ_e	2.34	rate exponential term, workers
η	0.85	span of control	σ_z	0.48	std. dev. Gaussian term, entrep.
ε/σ	0.15	super-elasticity of demand	λ_z	3.13	rate exponential term, entrep.
ω	0.12	fraction entrepreneurs	σ	8.81	demand elasticity at $q = 1$

Table 2: Moments Used in Calibration of Economy with Markups

	Data	Model
Income share, entrepreneurs	0.32	0.31
Gini income, all	0.64	0.64
Gini income, entrepreneurs	0.68	0.68
Gini income, workers	0.58	0.58
Income share top 1%, all	0.21	0.22
Income share top 1%, entrepreneurs	0.24	0.23
Income share top 1%, workers	0.14	0.13

labor shares to sales of -3.1%, a number reported by Edmond et al. (2023).[4] To understand why we adopt a different calibration strategy, recall that, as discussed in the main text, the Klenow and Willis (2016) functional form assumption for the Kimball aggregator in the product market implies a direct link between a firm's markup and its market share that allows us to obtain an estimate of the super-elasticity ε/σ without solving the model. Absent such a direct relationship in the economy with markdowns, we adopt an indirect inference approach and target instead the rate at which the labor share decreases with firm size. We note however that our baseline model, as well as the economy with markups only, also closely reproduce this statistic even though we have not explicitly targeted it. Table 3 reports the parameter values and Table 4 reports the model fit. These parameter values imply an aggregate markdown of 0.81.

Figure 1 illustrates the workings of the economy with markdowns. More productive firms hire more labor, so their relative employment $x(z)$ is larger, implying a lower markdown, as shown in the left panel of the figure. Because more productive firms pay their workers too little, they hire too few workers, which once again distorts production efficiency. As the right panel of the figure shows, a planner who seeks to maximize the total amount of output produced given an aggregate labor input l would prescribe that more productive firms expand

[4] Edmond et al. (2023) report an elasticity of markups with respect to sales equal to 3.1%. Since they calculate markups as the inverse of the labor share, implicitly assuming no markdown distortions that systematically vary with firm sales, their estimates imply an elasticity of the labor share to firm sales equal to -3.1%.

Table 3: Parameter Values in Economy with Markdowns

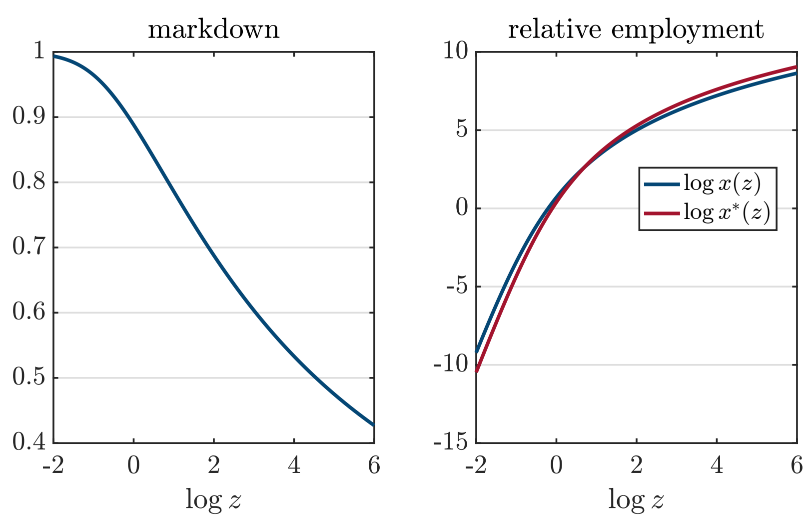
Assigned			Calibrated		
θ	1	CRRA coefficient	σ_e	0.99	std. dev. Gaussian term, workers
γ	2	inverse Frisch elasticity	λ_e	2.27	rate exponential term, workers
η	0.85	span of control	σ_z	0.52	std. dev. Gaussian term, entrep.
ω	0.12	fraction entrepreneurs	λ_z	4.47	rate exponential term, entrep.
			ν	9.80	labor elasticity at $x = 1$
			ϵ/ν	0.30	super-elasticity of labor

Table 4: Moments Used in Calibration of Economy with Markdowns

	Data	Model
Income share, entrepreneurs	0.31	0.31
Gini income, all	0.63	0.64
Gini income, entrepreneurs	0.69	0.68
Gini income, workers	0.58	0.58
Income share top 1%, all	0.20	0.22
Income share top 1%, entrepreneurs	0.24	0.23
Income share top 1%, workers	0.14	0.13
Elasticity labor-share to firm-size, $\times 100$	-3.1	-3.3

at the expense of smaller firms.

Figure 1: Markdowns and Firm Productivity



Because in this economy, as in the economy with markups, larger firms enjoy more market power, an increase in product market concentration caused by an increase in the dispersion of firm productivity which leaves the first-best level of aggregate productivity unchanged, reduces the aggregate markdown, aggregate productivity and the equilibrium wage. To illustrate this, consider the effect of doubling the variance of firm productivity. This increases the sales share of the largest 10% of firms from 0.52 to 0.75, increases the losses from misallocation from 0.44% to 1.06%, reduces the aggregate markdown from 0.81 to 0.76 and consequently reduces the equilibrium wage by 6.8%. Thus, as in the economy with markups, more product market concentration, in and of itself, reduces equilibrium wages and the welfare of workers.

2.3.2 Optimal Policy Experiments

Figure 2 shows the optimal wedges introduced by a utilitarian regulator in the economies with markups and markdowns, respectively. As in the economy with markups, concerns for redistribution imply a non-constant wedge in the economy with markdowns, so the regulator does not restore allocative efficiency. Nevertheless, the wedges chosen by the regulator are less steep than the wedges in the status quo allocations, suggesting that the regulator finds it optimal to reallocate employment towards more productive entrepreneurs, improving allocative efficiency.

Figure 2: Optimal Wedges: Utilitarian Regulator

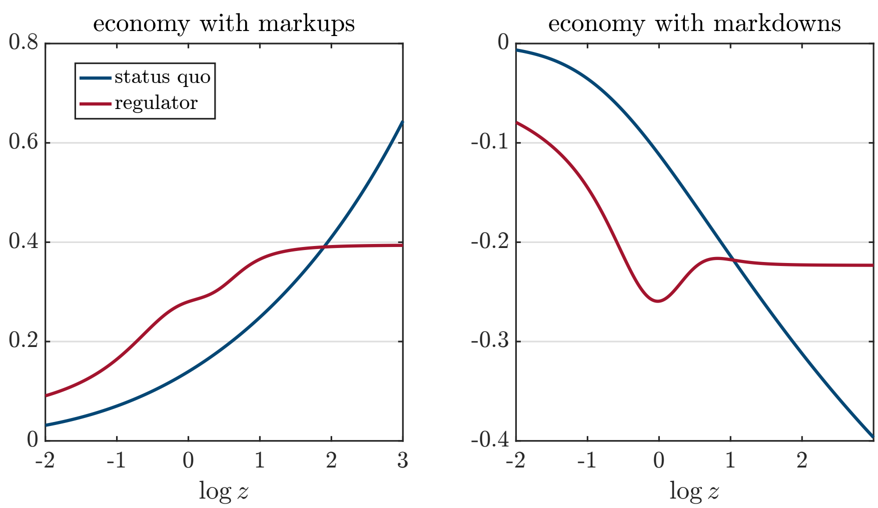


Table 5 reports the product market concentration implied by optimal regulation in the two economies. We contrast it with the product market concentration in the status quo and with that required to restore allocative efficiency. Because both the markup and markdown wedges change with firm size at a similar rate in the two economies, the degree of product market concentration required to restore allocative efficiency increases by a similar amount relative to the status quo of the two economies. For example, the sales share of the largest 5% of producers increases from 0.40 to 0.45 in the economy with markups and from 0.39 to 0.44 in the economy with markdowns. A utilitarian regulator nearly restores allocative efficiency in both economies: the sales share of the largest 5% of firms increases to 0.43 in the economy with markups and 0.45 in the economy with markdowns. As in our baseline model, a regulator who only values the welfare of workers prescribes even higher product market concentration: the sales share of the largest 5% of firms increases to 0.48 in the economy with markups and 0.50 in the economy with markdowns. We thus conclude that the policy implications we derived in the baseline economy are robust to considering alternative sources of firm market power.

Table 5: Product Market Concentration in Economies with Markups and Markdowns Only

	Status quo	Efficient allocation	Optimal regulation utilitarian	Optimal regulation only workers
Panel A. Economy with Markups Only				
Sales share top 1%	0.19	0.23	0.22	0.24
Sales share top 5%	0.40	0.45	0.43	0.48
Sales share top 10%	0.52	0.58	0.55	0.62
Panel B. Economy with Markdowns Only				
Sales share top 1%	0.19	0.23	0.23	0.26
Sales share top 5%	0.39	0.44	0.45	0.50
Sales share top 10%	0.52	0.57	0.58	0.65

3 Survey of Consumer Finances

The Survey of Consumer Finances (SCF) is a survey conducted by the National Opinion Research Center at the University of Chicago. This survey is well suited for characterizing the earnings, income, and wealth concentration at the top because it over-samples rich households. The unit of observation we use is the household. Each wave of the survey samples more than 6,000 households and is representative of the US economy.

Sample Selection. As is standard in the literature, we exclude households with negative income. In addition, we focus on a sample of households in which the household head is between 22 and 79 years old.

Wealth. Our measure of household wealth is the variable constructed by the Federal Reserve for its Bulletin article which accompanies each wave of the SCF. Wealth is defined as total net worth, which equals assets minus debt. Assets include both financial and non-financial assets. Financial assets include checking and savings accounts, stocks held directly and indirectly, bonds, etc. Non-financial assets, among others, include the value of houses and other real estate, the value of farms and private businesses owned by the household.[5] Debt includes both housing debt (e.g. mortgages), debt from lines of credit or credit cards, installment loans, etc.

Income. Our measure of income includes all sources of income excluding government transfers (e.g. social security and unemployment benefits) and excluding other (non-classified) sources of income. Thus, we include wage income, income from businesses, income from interests and dividends, from capital gains, rent income and income from pensions and annuities.

Definition of entrepreneurs. In contrast to Cagetti and De Nardi (2006), we consider a broader measure of entrepreneurship that includes all households in which the household head owns a business, excluding those who own C-corporations.[6]

4 Robustness Economies

4.1 Static Model

Figure 3 shows how the optimal degree of product market concentration, measured by the sales share of the largest 5% producers, changes as we vary key model parameters. In the top-left panel, we vary the super-elasticity ε/σ, adjusting the value of σ to ensure that the aggregate markup is unchanged. A utilitarian regulator increases product market concentration relative to the status quo by a larger amount when the super-elasticity is larger and markups increase faster with firm market shares. Only when ε/σ is low enough, less than 0.1, does the regulator reduce product market concentration. This is because in this region the markup distortion is low, so the desire to redistribute to poor entrepreneurs dominates the efficiency concerns. In contrast, when the regulator only values the welfare of workers, it always chooses to increase product market concentration relative to the status quo, as well as relative to what is required to restore allocative efficiency. This is because encouraging firms to expand bids up the labor share and the equilibrium wage.

In the top-right panel we vary σ, holding ε/σ unchanged, and thus tracing out the effect of changing the level of the aggregate markup. The lower σ is, and thus the higher the aggregate markup, the higher is optimal product market concentration relative to the status quo. As in the previous experiment, when σ is high and therefore the markup distortion is

[5] The value of houses, real estate and businesses is self-reported. E.g. with respect to housing the survey asks: "What is the current value of this (home and land/apartment/property)?". For businesses, the survey asks: "What is the net worth of (your share of) this business?"

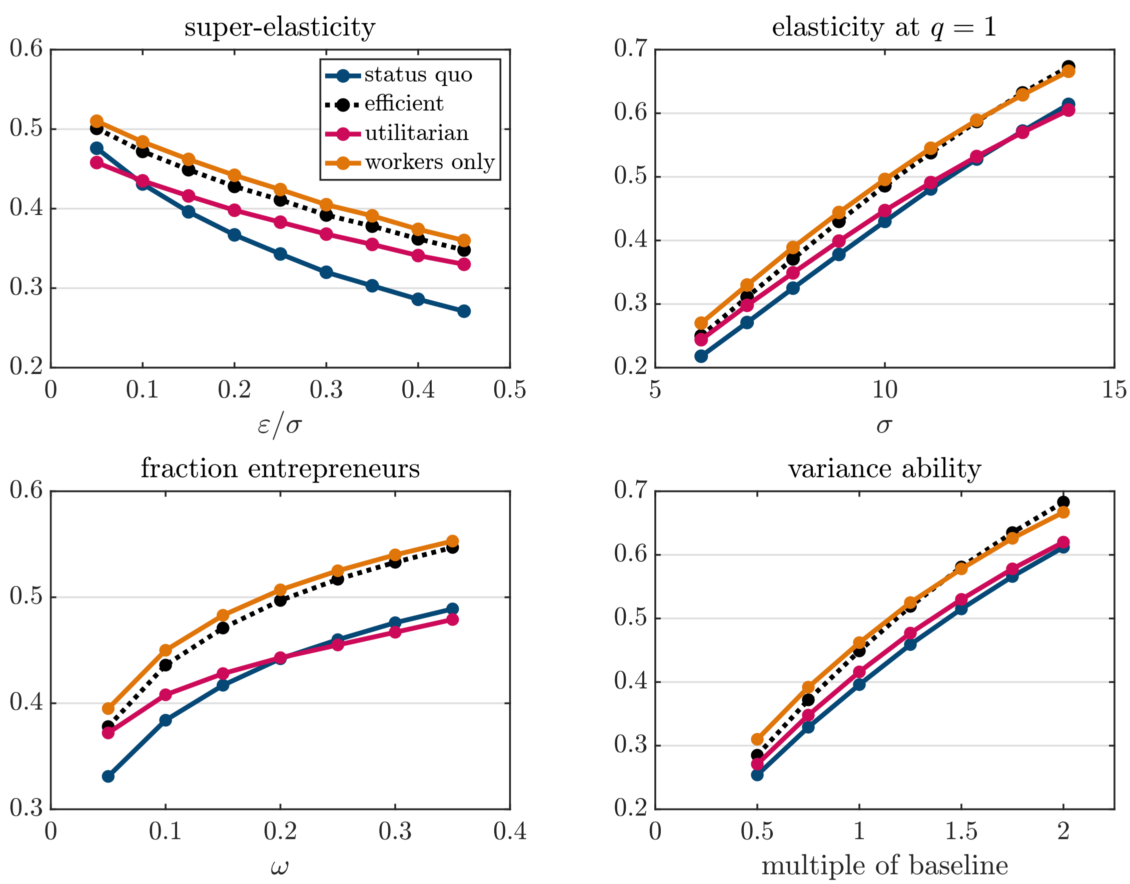
[6] The exact question in the survey is: (does the household head) "own privately-held businesses?" The SCF reports the legal status of up to two businesses own by the household. We identify households as owners of C-corporations if at least one of their businesses is reported to be a C-corporation.

low, a utilitarian regulator reduces concentration relative to the status quo. Once again, the regulator who only maximizes the welfare of workers increases product market concentration relative to the status quo.

In the bottom-left panel, we vary the share of entrepreneurs in the population, ω. For values of ω less than 20%, the empirically relevant range, optimal regulation leads to more product market concentration, more so when the regulator only values the welfare of workers. When the share of entrepreneurs is sufficiently high, a utilitarian regulator reduces product market concentration in an effort to redistribute to the less productive entrepreneurs. Intuitively, as ω increases, this motive becomes stronger and dominates the desire to increase the labor share. A regulator who only seeks to maximize the welfare of workers does not have this motive, so it prescribes more product market concentration regardless of the share of entrepreneurs in the economy.

In the bottom-right panel we scale the variance of labor and entrepreneurial ability relative to the benchmark calibration and find that, once again, optimal regulation implies more product market concentration than under the status quo, especially when the regulator seeks to only maximize the welfare of workers.

Figure 3: Top 5% Sales Share



4.2 Dynamic Model

In Section 3.8 we showed that our results are robust to alternative values of the super-elasticity ε/σ, the share of entrepreneurs ω, the skill-bias ζ, the relative risk aversion θ, as well as to an alternative parameterization of the processes for entrepreneurial and labor market ability that allows for a fat-tailed distribution of ability to better match top income and wealth inequality. Here we provide more details about these experiments. In particular, we report the parameter values and moments in each of these economies. Unless otherwise noted, we recalibrate each of these economies to match the same set of moments as in our baseline model.

Table 6 summarizes the parameterization of the alternative economy in which we set $\varepsilon/\sigma = 0.3$, a value twice as large as in the benchmark. All other assigned parameters are as in the benchmark and we recalibrate the remaining parameters to match the same set of targets as in the benchmark.

Table 6: Economy with Higher ε/σ

Panel A. Parameter Values

Assigned			Calibrated		
θ	1	CRRA	β	0.966	discount factor
γ	2	inverse Frisch	ρ_e	0.992	AR(1) e
α	1/3	capital elasticity	σ_e	0.128	std. dev. e shocks
η	0.85	span of control	ρ_z	0.989	AR(1) z
δ	0.06	capital depreciation rate	σ_z	0.094	std. dev. z shocks
ρ	1.41	skill elasticity of substitution	σ	14.68	demand elasticity at $q = 1$
ω	0.12	fraction of entrepreneurs	μ_c	1.925	mean productivity corporations
ω_2	0.28	fraction high-skill workers	$\bar{\mathcal{K}}\bar{\vartheta}$	0.082	entry costs / GDP
ε/σ	0.30	super-elasticity of demand	$\bar{\psi}$	2.817	avg. elasticity of high-skill labor
φ	0.04	exit rate, corporations	ζ	0.719	sensitivity of elasticity high-skill labor

Panel B. Moments

	Data	Model		Data	Model
Wealth to income ratio	6.57	6.74	Gini wealth, workers	0.83	0.82
Wealth share of entrepr.	0.46	0.46	Gini income, workers	0.59	0.61
Income share of entrepr.	0.31	0.30	Skill premium	1.85	1.86
Gini wealth, all hhs	0.85	0.83	Wage-employment elasticity, $\times 100$	1.90	1.89
Gini income, all hhs	0.64	0.66	Fraction of corporate firms	0.05	0.05
Gini wealth, entrepr.	0.78	0.71	Sales share corporate firms	0.63	0.56
Gini income, entrepr.	0.68	0.72			

Table 7 reports the moments implied by a comparative statics experiment in which we increase the fraction of entrepreneurs, ω, to 0.2, leaving all other parameters unchanged.

Table 8 summarizes the parameterization of the alternative economy in which we target

Table 7: Moments in Economy with Higher ω

	Data	Model		Data	Model
Wealth to income ratio	6.57	6.56	Gini wealth, workers	0.83	0.81
Wealth share of entrepr.	0.46	0.45	Gini income, workers	0.59	0.59
Income share of entrepr.	0.31	0.26	Skill premium	1.85	1.97
Gini wealth, all hhs	0.85	0.81	Wage-employment elasticity, $\times 100$	1.90	1.88
Gini income, all hhs	0.64	0.64	Fraction of corporate firms	0.05	0.05
Gini wealth, entrepr.	0.78	0.75	Sales share corporate firms	0.63	0.66
Gini income, entrepr.	0.68	0.75			

a wage-employment elasticity of 3.8%, a value twice as large as in the benchmark and Bloom et al. (2018). All other assigned parameters are as in the benchmark and we recalibrate the remaining parameters to match the same set of remaining targets as in the benchmark.

Table 8: Economy with Higher ζ

Panel A. Parameter Values

Assigned			Calibrated		
θ	1	CRRA	β	0.966	discount factor
γ	2	inverse Frisch	ρ_e	0.992	AR(1) e
α	1/3	capital elasticity	σ_e	0.129	std. dev. e shocks
η	0.85	span of control	ρ_z	0.987	AR(1) z
δ	0.06	capital depreciation rate	σ_z	0.093	std. dev. z shocks
ρ	1.41	skill elasticity of substitution	σ	7.551	demand elasticity at $q = 1$
ω	0.12	fraction of entrepreneurs	μ_c	1.839	mean productivity corporations
ω_2	0.28	fraction high-skill workers	$\bar{\mathcal{K}}\bar{\vartheta}$	0.072	entry costs / GDP
ε/σ	0.15	super-elasticity of demand	$\bar{\psi}$	4.121	avg. elasticity of high-skill labor
φ	0.04	exit rate, corporations	ζ	1.056	sensitivity of elasticity high-skill labor

Panel B. Moments

	Data	Model		Data	Model
Wealth to income ratio	6.57	6.31	Gini wealth, workers	0.83	0.81
Wealth share of entrepr.	0.46	0.41	Gini income, workers	0.59	0.59
Income share of entrepr.	0.31	0.30	Skill premium	1.85	1.94
Gini wealth, all hhs	0.85	0.81	Wage-employment elasticity, $\times 100$	3.80	3.72
Gini income, all hhs	0.64	0.64	Fraction of corporate firms	0.05	0.05
Gini wealth, entrepr.	0.78	0.68	Sales share corporate firms	0.63	0.54
Gini income, entrepr.	0.68	0.64			

Table 9 summarizes the parameterization of the alternative economy in which we set the relative risk aversion $\theta = 2$, a value twice as large as in the benchmark. All other assigned parameters are as in the benchmark and we recalibrate the remaining parameters to match the same set of targets as in the benchmark.

Table 9: Economy with Higher θ

Panel A. Parameter Values

Assigned			Calibrated		
θ	2	CRRA	β	0.958	discount factor
γ	2	inverse Frisch	ρ_e	0.993	AR(1) e
α	1/3	capital elasticity	σ_e	0.146	std. dev. e shocks
η	0.85	span of control	ρ_z	0.988	AR(1) z
δ	0.06	capital depreciation rate	σ_z	0.102	std. dev. z shocks
ρ	1.41	skill elasticity of substitution	σ	9.263	demand elasticity at $q = 1$
ω	0.12	fraction of entrepreneurs	μ_c	1.699	mean productivity corporations
ω_2	0.28	fraction high-skill workers	$\bar{\mathcal{K}}\bar{\vartheta}$	0.075	entry costs / GDP
ε/σ	0.15	super-elasticity of demand	$\bar{\psi}$	2.615	avg. elasticity of high-skill labor
φ	0.04	exit rate, corporations	ζ	0.606	sensitivity of elasticity high-skill labor

Panel B. Moments

	Data	Model		Data	Model
Wealth to income ratio	6.57	6.60	Gini wealth, workers	0.83	0.83
Wealth share of entrepr.	0.46	0.47	Gini income, workers	0.59	0.59
Income share of entrepr.	0.31	0.28	Skill premium	1.85	1.83
Gini wealth, all hhs	0.85	0.85	Wage-employment elasticity, $\times 100$	1.90	1.91
Gini income, all hhs	0.64	0.65	Fraction of corporate firms	0.05	0.05
Gini wealth, entrepr.	0.78	0.73	Sales share corporate firms	0.63	0.62
Gini income, entrepr.	0.68	0.73			

We next consider an extension with a super-star state that allows the model to match the top income and wealth shares, as in Castaneda et al. (2003). We assume that an agent can be in either a normal or a super-star state. In the normal state labor market ability follows an AR(1) process as earlier. In the super-star state, labor market ability is relatively high, $\bar{e}$ times higher than the average. We assume that agents transit from the normal to the super-star state with probability p_e and remain in the super-star state with probability q_e. When agents return to the normal state, they draw a new ability level from the ergodic distribution associated with the AR(1) process. An analogous process for entrepreneurial ability is characterized by parameters ρ_z, σ_z, p_z, q_z and $\bar{z}$.

To calibrate the additional parameters describing the super-star state, we augment the original set of moments we target with statistics describing the wealth and income shares of the top 1% of households, as well as the top 1% of workers and entrepreneurs in isolation. Table 10 reports the calibrated parameter values in this economy. Table 11 shows that the model reproduces the targeted moments well. For example, the top 1% of households hold 35% of all wealth in the data, 37% in this calibration, and 28% in our baseline model without a super-star state.

Table 10: Parameter Values in Economy with Super-Star State

β	0.966	discount factor
ρ_e	0.993	AR(1) e
σ_e	0.118	std. dev. e shocks
p_e	1.5e-6	prob. enter super-star e state
q_e	0.977	prob. stay super-star e state
$\bar{e}$	6.449	log ability super-star e state, rel. mean
ρ_z	0.989	AR(1) z
σ_z	0.099	std. dev. z shocks
p_z	1.2e-5	prob. enter super-star z state
q_z	0.950	prob. stay super-star z state
$\bar{z}$	3.535	log ability super-star z state, rel. mean
σ	8.601	demand elasticity at $q = 1$
μ_c	1.798	mean productivity corporations
$\bar{\mathcal{K}}$	0.077	fixed entry cost / GDP
$\bar{\psi}$	3.060	avg. elasticity of high-skill
ζ	0.634	sensitivity of elasticity high-skill labor

Table 11: Moments Used to Calibrate Economy with Super-Star State

	Data	Model		Data	Model
Wealth to income ratio	6.57	6.59	Wealth share top 1%	0.35	0.35
Wealth share of entrepr.	0.46	0.47	Income share top 1%	0.22	0.22
Income share of entrepr.	0.31	0.29	Wealth share top 1% entrepr.	0.24	0.23
Gini wealth, all hhs	0.85	0.87	Income share top 1% entrepr.	0.23	0.23
Gini income, all hhs	0.64	0.67	Wealth share top 1% workers	0.31	0.30
Gini wealth, entrepr.	0.78	0.76	Income share top 1% workers	0.16	0.15
Gini income, entrepr.	0.68	0.75	Skill premium	1.85	1.84
Gini wealth, workers	0.83	0.86	Wage-employment elasticity, $\times 100$	1.90	1.88
Gini income, workers	0.59	0.61	Fraction of corporate firms	0.05	0.05
			Sales share corporate firms	0.63	0.61

References

Berger, David, Kyle Herkenhoff, and Simon Mongey, "Labor Market Power," *American Economic Review*, April 2022, *112* (4), 1147–93.

Bloom, Nicholas, Fatih Guvenen, Benjamin S. Smith, Jae Song, and Till von Wachter, "The Disappearing Large-Firm Wage Premium," *AEA Papers and Proceedings*, May 2018, *108*, 317–22.

Boar, Corina and Virgiliu Midrigan, "Markups and Inequality," Working Paper 2022.

Cagetti, Marco and Mariacristina De Nardi, "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, 2006, *114* (5), 835–870.

Castaneda, Ana, Javier Diaz-Gimenez, and Jose-Victor Rios-Rull, "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 2003, *111* (4), 818–857.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, “How Costly Are Markups?,” *Journal of Political Economy*, 2023, *131* (7), 1619–1675.

Klenow, Peter J. and Jonathan L. Willis, “Real Rigidities and Nominal Price Changes,” *Economica*, July 2016, *83*, 443–472.