Markups and Inequality*

Corina Boar[†] Virgiliu Midrigan[‡]

April 2024

Abstract

We characterize optimal product market policy in an unequal economy in which firm ownership is concentrated and markups increase with firm market shares. We study the problem of a utilitarian regulator who designs revenue-neutral interventions in the product market. We show that optimal policy increases product market concentration. This is because policies that encourage larger producers to expand improve allocative efficiency, increase the demand for labor and equilibrium wages. We derive these results both in a static Mirrleesian setting in which we impose no constraints on the shape of interventions, as well as in a dynamic economy with wealth accumulation. In our dynamic economy optimal policy reduces wealth and income inequality by redistributing market share and profits from medium-sized businesses, which are primarily owned by relatively rich entrepreneurs, to larger diversified corporate firms.

Keywords: entrepreneurs, inequality, markups, misallocation, redistribution.

JEL classifications: D4, E2, L1.

*We thank the editor, Nir Jaimovich, and five anonynmous referees for valuable feedback. We also thank Federico Kochen for superb research assistance, our discussants Salome Baslandze, Paco Buera, Basile Grassi and Aleh Tsyvinski, as well as Adrien Auclert, Jess Benhabib, Mariacristina De Nardi, Mark Gertler, Mike Golosov, Oleg Itskhoki, Benjamin Moll, Gianluca Violante, Ludwig Straub and seminar participants at Barcelona GSE Forum, Boston University, Cleveland Fed, Copenhagen, Duke, Dutch National Bank Conference on Inequality, ERMAS, Harvard, Ischia Conference on Frontiers of Macroeconomics, Madrid Conference on Firm Dynamics and Macroeconomics, Minneapolis Fed, National Bank of Romania, NBER Summer Institute, NYU, OSU, Penn, Philadelphia Fed, Princeton, Rochester, Rome Conference on Macroeconomics, Sciences Po, SED, SITE, St Louis Conference on Macroeconomics of Inequality, Stony Brook and Warwick Conference on Inequality for valuable comments. We gratefully acknowledge support from the National Science Foundation, Grant SES-1948119.

†New York University and NBER, corina.boar@nyu.edu.

‡New York University and NBER, virgiliu.midrigan@nyu.edu.

1 Introduction

The United States has experienced a sharp increase in product market concentration, profits and measured markups in recent decades, in large part due to the rise of superstar firms.[1] Since firm ownership is highly concentrated, a growing concern is that markups redistribute income from workers toward firm owners, thus increasing inequality. This led to numerous calls for rethinking competition policy to explicitly incorporate distributional concerns, in addition to concerns for economic efficiency.[2]

Existing work on markups, such as Atkeson and Burstein (2008), Bilbiie et al. (2012, 2019) and Edmond et al. (2023), assumes perfect consumption sharing and thus abstracts from distributional considerations. In such a setting markups only distort production by introducing two sources of inefficiency. First, the aggregate markup acts as a uniform tax on production. Second, firms with higher market shares charge higher markups, and the resulting dispersion in marginal products reduces allocative efficiency and aggregate productivity. In this environment a policy that subsidizes production in proportion to markups restores efficiency. Even though this policy increases concentration and profits, it makes the representative consumer, who owns all firms, better off. This policy prescription ignores, however, the tradeoff between equity and efficiency that arises in an unequal economy.

Our paper departs from the representative consumer framework. We study optimal product market policy in an economy that matches the degree of inequality in the United States and in which firm ownership is highly concentrated and markups increase with firm market shares. In addition, we assume assortative matching between firms and workers, so that more productive firms disproportionately hire high-skill workers. In this economy, an increase in product market concentration caused by higher dispersion in firm productivity raises the aggregate markup, reduces the labor share, and increases the skill premium, thus exacerbating income inequality.

Our main finding is that optimal policy encourages larger producers to expand, therefore increasing product market concentration. Even though this policy raises markups and the skill premium, it improves allocative efficiency and bids up the demand for labor, thus increasing the labor share and wages. The higher the markup distortion in the initial economy and the higher the welfare weight that the policy maker places on workers relative to firm owners, the higher the product market concentration that optimal policy prescribes.

[1] De Loecker et al. (2020), Hall (2018a), Autor et al. (2020), Hartman-Glaser et al. (2019).

[2] Stiglitz (2012), Atkinson (2015), Baker and Salop (2015) and Khan and Vaheesan (2016).

We develop our argument in two steps. In the first part of the paper, we build intuition by studying a static economy in which we use a mechanism design approach to characterize optimal product market interventions. We show that the optimal allocations can be implemented using size-dependent production subsidies and taxes. Moreover, a simple parametric subsidy function can achieve the bulk of the gains from unrestricted regulation. In the second part of the paper, we conduct a quantitative analysis in a richer dynamic economy in which private businesses compete alongside corporate firms. For computational tractability, here we restrict attention to policies in the simple parametric class.

The static economy we study consists of workers and entrepreneurs. There are two types of workers, low- and high-skill, who are heterogeneous in their labor market efficiency and choose how many hours to work. Entrepreneurs differ in their ability, hire labor, and supply a differentiated variety of a good. The assumptions we make on the demand system imply that the demand elasticity a producer faces decreases in its market share, so larger producers charge higher markups. The assumptions we make on technology imply assortative matching between firms and workers: higher productivity firms disproportionately hire high-skill workers. Our framework thus parsimoniously captures the trade-off between efficiency gains and markups that is at the heart of the debate about product market policies.

We build on the approach of Baron and Myerson (1982) who study the problem of regulating a single monopolist. In contrast to their work, we consider the problem of regulating all firms in a general equilibrium setting. We assume that the regulator does not observe the ability of individual entrepreneurs and thus faces incentive compatibility constraints. These constraints generate informational rents, which increase with the equilibrium wages and the amount of output the regulator prescribes that the entrepreneur produces.

We characterize the optimal product market interventions of a utilitarian regulator, restricting attention to *revenue-neutral* interventions. The regulator can shape the firm size and markup distribution, but cannot raise revenue from firms to fund direct transfers to consumers.[3] The regulator thus recognizes that it can only increase the welfare of workers indirectly by increasing the equilibrium wages and balances the following tradeoff between equity and efficiency. Reducing the market share of productive entrepreneurs allows the reg-

[3] As we showed in Boar and Midrigan (2019), absent this restriction, the regulator would still find it optimal to increase product market concentration by taxing smaller firms more than larger ones but would increase the average output tax, thus raising revenue to finance lump-sum transfers. Such a policy can be mimicked by higher income taxes, the study of which is beyond the scope of this paper. See da Costa and Maestri (2019), Kaplow (2021), Kushnir and Zubrickas (2019), Jaravel and Olivi (2021) and Eeckhout et al. (2021) who study income taxation in economies with imperfectly competitive markets.

ulator to redistribute to less productive entrepreneurs and workers on one hand, but on the other hand, it reduces productivity and wages.

A robust result that emerges is that optimal regulation leads to a higher degree of product market concentration relative to the status quo. Though optimal interventions do not fully restore allocative efficiency, the degree of product market concentration is nearly as large as that implied by the efficient allocations. Perhaps counter-intuitively, product market concentration is higher when the regulator places a higher weight on the welfare of workers. This is because product market interventions that encourage larger firms to expand bid up the demand for labor and therefore the equilibrium wages.

We show that one can implement the optimal policy with an output subsidy schedule. Though this schedule is highly non-linear, it can be well approximated by a simple three-parameter subsidy function.[4] These parameters determine the lump-sum transfer to individual producers, the average marginal subsidy, and the slope of the marginal subsidy schedule, thus allowing us to provide a sharper intuition for the tradeoffs the regulator faces. We show that increasing the slope of the marginal subsidy schedule increases product market concentration and income inequality, but leads to higher productivity and a higher labor share. Though high-skill workers disproportionately benefit from steeper marginal output subsidies, low-skill workers benefit as well, despite the increase in the skill premium.

Our static model is purposefully simple in order to highlight the key tradeoffs between equity and efficiency entailed by product market interventions. We show, however, that our conclusions extend to a richer dynamic setting in which we introduce capital and wealth accumulation, a corporate sector whose ownership is diversified, and a government that provides some redistribution via income taxes and transfers. We restrict product market interventions to the three-parameter subsidy class and study optimal regulation explicitly taking into account that product market reforms generate long-lasting transition dynamics. As in the static model, we find that optimal intervention increases the market share of the largest firms, especially when the regulator is only concerned with maximizing the welfare of workers. Though the welfare gains from optimal product market interventions are modest relative to what can be achieved with direct redistribution through income taxes (as in, for example, Boar and Midrigan, 2022a), suggesting that competition policy is too blunt a tool to address inequality, our results imply that efficiency and equity motives are aligned because optimal interventions nearly restore allocative efficiency.

[4] See Heathcote and Tsujiyama (2021) for an analogous exercise in the context of labor income taxation.

Our result that optimal policy encourages larger firms to expand is robust to perturbations of the key parameters of the model, including the super-elasticity of demand which determines how rapidly markups increase with firm market shares, the share of entrepreneurs, the degree of assortative matching between firms and workers, the coefficient of relative risk aversion, and the process for entrepreneurial and labor market ability. In the Appendix we also study a version of our model in which firms have labor market power and show that optimal policy once again implies an increase in product market concentration.

We conclude that product market concentration is not necessarily costly, even in an environment with highly unequal firm ownership. What is costly is dispersion in the marginal product of factors of production across firms and wedges that depress the equilibrium wages and the return on capital. Optimal product market interventions reduce these wedges, even though they increase product market concentration. Our results thus caution against the widely-held view that reducing concentration and the market power of large firms necessarily improves the welfare of the poor. Though less concentration indeed reduces market power and markups in our model, the interventions required to reduce the market share of large firms have the unintended consequence of also reducing the labor share, aggregate productivity and equilibrium wages.

Related Work. In addition to the work on markups and optimal taxation discussed above, our paper builds on studies of wealth and income inequality, originating with Castaneda et al. (2003) and more recently Benhabib et al. (2017) and Hubmer et al. (2021). This work typically assumes perfect competition in the product market or constant markups. Several notable exceptions are Brun and Gonzalez (2017) and Colciago and Mechelli (2019) who study the effect of increasing markups in Bewley-Aiyagari models with homogeneous firms. In contrast to their work, we explicitly model firm heterogeneity and study optimal product market interventions. A recent paper by Dworczak et al. (2021) also considers a mechanism design approach to characterize the tradeoff between efficiency and redistribution in a setting in which buyers and sellers differ in their valuation of a good. In contrast to their paper, which studies a market for a single good, we study a production economy with a large number of goods and account for the general equilibrium effects of regulation.

Our paper is also related to a large literature on markups, product market misallocation and size-dependent policies (Guner et al., 2008, Restuccia and Rogerson, 2008, Hsieh and Klenow, 2009, Jones, 2011, Baqaee and Farhi, 2019, Pellegrino, 2024). We show that in our economy, concerns for inequality prevent optimal product market interventions from fully eliminating misallocation.

The remainder of the paper proceeds as follows. Section 2 describes the static economy we study and solves the optimal regulation problem. Section 3 extends the analysis to a dynamic setting. Section 4 concludes.

2 Static Model

For clarity, we study a simple environment that captures the interplay between markups and inequality and allows us to highlight the key forces that shape optimal product market interventions and motivate the policy experiments we conduct in the richer dynamic model we study in Section 3.

2.1 Environment

The economy is inhabited by a measure ω of entrepreneurs and a measure $1-\omega$ of workers. A measure ω_1 of these workers are low-skill and a measure $\omega_2 = 1-\omega-\omega_1$ are high-skill. Workers are heterogeneous in their labor market ability e and choose how much to work at a wage W_s, $s \in \{1, 2\}$. Entrepreneurs are heterogeneous in their entrepreneurial ability z. They hire labor, supply a differentiated variety of a good and receive income from profits. We first describe the problem of the agents, characterize the equilibrium in the absence of product market interventions, and discuss the distortions due to markups.

2.1.1 Workers

All workers have preferences of the form

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

where c denotes consumption and h hours worked. The parameters θ and γ represent the coefficient of relative risk aversion and the inverse of the Frisch elasticity of labor supply. The budget constraint of workers of type s is

$$c = W_s e h.$$

Solving workers' problems gives their optimal hours and consumption choices

$$h_s(e,W_s) = (W_s e)^{\frac{1-\theta}{\gamma+\theta}} \qquad \text{and} \qquad c_s(e,W_s) = (W_s e)^{\frac{1+\gamma}{\gamma+\theta}}\,. \tag{1}$$

The welfare of workers thus increases with the equilibrium wages

$$v_s(e, W_s) = u\left(c_s\left(e, W_s\right),\ h_s\left(e, W_s\right)\right) = \frac{\gamma+\theta}{\left(1-\theta\right)\left(1+\gamma\right)} W_s^{\frac{(1-\theta)(1+\gamma)}{\gamma+\theta}} e^{\frac{(1-\theta)(1+\gamma)}{\gamma+\theta}}. \tag{2}$$

2.1.2 Entrepreneurs

Entrepreneurs do not work and have preferences of the form

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

They differ in their ability z and operate a production technology

$$y = z\left[\left(1-\psi\left(z\right)\right)^{\frac{1}{\rho}} l_1^{\frac{\rho-1}{\rho}} + \psi\left(z\right)^{\frac{1}{\rho}} l_2^{\frac{\rho-1}{\rho}}\right]^{\eta\frac{\rho}{\rho-1}}, \tag{3}$$

where y is output, l_1 and l_2 are the efficiency units of labor of each skill type, ρ is the elasticity of substitution between the two skill types, $\eta \leq 1$ is the span-of-control parameter and $\psi\left(z\right) \in (0,1)$ governs the elasticity of high-skill labor in production.[5] We assume

$$\psi\left(z\right) = \frac{1}{1+\bar{\psi}z^{-\zeta}}, \tag{4}$$

where ζ determines the extent to which the relative demand for high-skill labor varies with firm productivity and $\bar{\psi}$ determines the average demand for high-skill labor and, therefore, the skill premium W_2/W_1. If $\zeta > 0$, then there is assortative matching between firms and workers.

The entrepreneur's budget constraint is

$$c = \pi = p(y)y - W_1 l_1 - W_2 l_2,$$

where π are profits and $p(y)$ is the inverse demand function, which we derive next.

Market Structure. A perfectly competitive final good sector aggregates differentiated varieties produced by entrepreneurs. Each entrepreneur is the only supplier of a given variety. The technology of the final good sector is implicitly defined by the Kimball aggregator

$$\int_0^{\omega} \Upsilon\left(\frac{y_i}{Y}\right) \mathrm{d}i = 1, \tag{5}$$

[5] See Burstein and Vogel (2017) and Waugh (2018), who use a similar production function to study the role of trade and immigration in shaping the skill premium.

where Y is the output of the final good, whose price we normalize to 1. The inverse demand function faced by an entrepreneur i is

$$p\left(y_i\right) = \Upsilon'\left(\frac{y_i}{Y}\right) D, \tag{6}$$

where

$$D = \left(\int_0^\omega \Upsilon'\left(\frac{y_i}{Y}\right)\frac{y_i}{Y}\,\mathrm{d}i\right)^{-1}$$

is an endogenously determined demand index.

We use the Klenow and Willis (2016) functional form for the aggregator $\Upsilon(q)$, which implies a demand elasticity

$$-\frac{\Upsilon'(q)}{\Upsilon''(q)q} = \sigma q^{-\frac{\varepsilon}{\sigma}},$$

that falls with the entrepreneur's relative quantity $q = y/Y$ or, equivalently, market share and implies that markups increase with firm market shares.[6] We note that such a demand system can be micro-founded by explicitly modeling consumer search frictions (Benabou, 1988) and that models of oligopolistic competition (Atkeson and Burstein, 2008) give rise to a similar positive relationship between market shares and markups.

Optimal Choices. The entrepreneur chooses its price, quantity and the labor input of each type to maximize profits. The optimal quantity choice is implicitly given by

$$D\Upsilon'\left(\frac{y\left(z\right)}{Y}\right) = m\left(z\right)\frac{1}{\eta}W\left(z\right)\left(\frac{y\left(z\right)}{z}\right)^{\frac{1}{\eta}}\frac{1}{y\left(z\right)}, \tag{7}$$

where the left-hand side is equal to the firm's price and the right-hand side is the product of the markup $m\left(z\right) = \frac{\sigma}{\sigma-\left(y(z)/Y\right)^{\frac{\varepsilon}{\sigma}}}$ and the marginal cost $\frac{1}{\eta}W\left(z\right)\left(\frac{y(z)}{z}\right)^{\frac{1}{\eta}}\frac{1}{y(z)}$. The term

$$W\left(z\right) = \left[\left(1-\psi\left(z\right)\right)W_1^{1-\rho} + \psi\left(z\right)W_2^{1-\rho}\right]^{\frac{1}{1-\rho}}$$

represents the composite wage index that an entrepreneur with productivity z faces.

The optimal choice of labor of each type is given by

$$W_s l_s\left(z\right) = \frac{\varepsilon_s\left(z\right)}{m\left(z\right)}p\left(z\right)y\left(z\right), \tag{8}$$

where

$$\varepsilon_2\left(z\right) = \eta\frac{\psi\left(z\right)^{\frac{1}{\rho}}l_2\left(z\right)^{\frac{\rho-1}{\rho}}}{\left(1-\psi\left(z\right)\right)^{\frac{1}{\rho}}l_1\left(z\right)^{\frac{\rho-1}{\rho}} + \psi\left(z\right)^{\frac{1}{\rho}}l_2\left(z\right)^{\frac{\rho-1}{\rho}}}$$

[6]See Chari et al. (2000), Gopinath and Itskhoki (2010) and Edmond et al. (2023), who also use the Kimball specification of the demand system.

is the elasticity of high-skill labor in production and $\varepsilon_1(z) = \eta - \varepsilon_2(z)$ is the elasticity of low-skill labor. Since $\psi'(z) > 0$, $\varepsilon_2'(z) > 0$ as well, so more productive entrepreneurs hire disproportionately more high-skill labor, implying positive assortative matching between firms and workers.

2.1.3 Aggregation

We assume that the distribution of labor market efficiency is the same across the two types of workers and is denoted by $H(e)$.[7] Aggregating their optimal choices in (1) gives the aggregate labor supply of type s

$$L_s^w(W_s) = \left(\omega_s \int_0^\infty e^{\frac{1+\gamma}{\gamma+\theta}} \, \mathrm{d}H(e)\right) W_s^{\frac{1-\theta}{\gamma+\theta}}$$

and the total consumption of workers of type s

$$C_s^w(W) = \left(\omega_s \int_0^\infty e^{\frac{1+\gamma}{\gamma+\theta}} \, \mathrm{d}H(e)\right) W_s^{\frac{1+\gamma}{\gamma+\theta}}.$$

Let $F(z)$ denote the distribution of entrepreneurial ability and $f(z)$ the corresponding density. Integrating the labor choices of individual entrepreneurs allows us to write the aggregate demand for labor of each type as

$$W_s L_s = \mathcal{E}_s \frac{Y}{\mathcal{M}}, \tag{9}$$

where

$$\mathcal{E}_s = \omega \int_0^\infty \varepsilon_s(z) \frac{W_1 l_1(z) + W_2 l_2(z)}{W_1 L_1 + W_2 L_2} \mathrm{d}F(z)$$

is the weighted average of the elasticity of type s labor in production, with weights given by the cost shares of individual entrepreneurs, and

$$\mathcal{M} = \omega \int_0^\infty m(z) \frac{W_1 l_1(z) + W_2 l_2(z)}{W_1 L_1 + W_2 L_2} \mathrm{d}F(z) \tag{10}$$

is the *aggregate markup*. This expression is identical to that derived by Edmond et al. (2023).

[7] In the quantification, we use education as a proxy for skills. As documented by Kuhn and Rios-Rull (2016), the Gini coefficient of earnings is similar across education groups. Since the average level of labor market efficiency is not pinned down separately from the price per efficiency unit of labor W_s, assuming it is the same across groups is without loss of generality.

2.2 Parameterization

Since the solution to this model is not attainable in closed form, we use numerical methods to solve for the optimal product market interventions. Even though, as we show in the Appendix, our conclusions are not driven by specific parameter choices, we find it useful to center our discussion around some empirically plausible parameter values.

We assign preference and technology parameters to values common in the literature. These are reported in the first column of Table 1. We set the share of entrepreneurs to $\omega = 0.117$, the fraction of respondents in the 2013 SCF who own a private pass-through business.[8] We associate high-skill workers in our model with workers with at least a college degree and set ω_2 to reproduce their population share in the 2013 CPS ASEC supplement. We assume an elasticity of substitution between high- and low-skill labor $\rho = 1.41$, as estimated by Katz and Murphy (1992).

Our choice of ε/σ follows Edmond et al. (2023), who use Census data to estimate the relationship between firm markups m_i and firm market share $s_i = p_i y_i / Y$ implied by the functional form of the Kimball aggregator[9]

$$\frac{1}{m_i} + \log\left(1 - \frac{1}{m_i}\right) = \text{const} + \frac{\varepsilon}{\sigma} s_i.$$

These researchers estimate a super-elasticity ε/σ in the neighborhood of 0.15. We use the same number here because we assume the same technology and thus our model implies the same relationship between markups and market shares. This value is also consistent with the estimates surveyed by Klenow and Willis (2016) and implies an elasticity of firm markups with respect to firm market shares of 2.9%, in line with the 3.1% estimate reported by Edmond et al. (2023).

We calibrate the remaining parameters, reported in the second column of Table 1, to match salient features of income inequality and assortative matching in the data. We follow Heathcote and Tsujiyama (2021) in assuming that the logarithm of idiosyncratic efficiency is drawn from an exponentially modified Gaussian distribution with parameters λ_i and σ_i for both workers ($i = e$) and entrepreneurs ($i = z$). Here σ_i represents the standard deviation of the Gaussian component and λ_i the rate coefficient of the exponential component. We choose these parameters to match the income share of entrepreneurs, the income Gini coefficients for

[8] A pass-through business is a business whose profits are taxed as personal income. See the Appendix for details.

[9] Edmond et al. (2023) estimate markups using a cost-share approach and define the market share as the firm's sales share in a 6-digit industry. See Appendix B of that paper for details.

Table 1: Parameter Values

Assigned			Calibrated		
θ	1	CRRA coefficient	σ_e	0.97	std. dev. Gaussian term, workers
γ	2	inverse Frisch elasticity	λ_e	2.37	rate exponential term, workers
η	0.85	span of control	σ_z	0.27	std. dev. Gaussian term, entrep.
ρ	1.41	skill elasticity of substitution	λ_z	3.21	rate exponential term, entrep.
ω	0.12	fraction entrepreneurs	σ	9.34	demand elasticity at $q = 1$
ω_2	0.28	fraction high-skill workers	$\bar{\psi}$	1.29	avg. elasticity of high-skill labor
ε/σ	0.15	super-elasticity of demand	ζ	0.37	sensitivity of elasticity high-skill labor

Table 2: Moments Used in Calibration

	Data	Model
Income share of entrepreneurs	0.31	0.31
Gini income, all households	0.64	0.64
Gini income, workers	0.59	0.58
Gini income, entrepreneurs	0.68	0.68
Income share top 1%, all households	0.22	0.20
Income share top 1%, workers	0.16	0.14
Income share top 1%, entrepreneurs	0.23	0.23
Skill premium	1.85	1.85
Wage-employment elasticity, ×100	1.90	1.90

all households, as well as for workers and entrepreneurs in isolation, and the income shares of the richest 1% of households in all sub-groups, as measured in the 2013 SCF. We target a skill premium of 1.85, as reported by Heathcote et al. (2023). We capture the extent of assortative matching by requiring that the model reproduces the 1.9% elasticity of average wages to firm employment documented by Bloom et al. (2018).[10] As Table 2 shows, the model matches the targeted moments well.

2.3 Markup Distortions and Implications for Inequality

We next discuss the sources of inefficiency introduced by markups. In this economy, markups generate two production distortions. First, as equation (9) shows, the *level* of the aggregate markup $\mathcal{M}$ acts as a uniform tax on both types of employment and depresses the equilibrium

[10] We focus on the component of earnings attributed to worker effects in the data, to isolate the role of assortative matching between firms and workers.

wages W_s relative to the marginal product of each type of labor $\mathcal{E}_s Y/L_s$. Our parameter choices imply that the aggregate markup is equal to 1.23, a number similar to the estimates of Edmond et al. (2023) and Hall (2018b) for 2013.

Second, as equation (8) shows, *dispersion* in markups generates dispersion in the marginal product of each type of labor. The resulting misallocation reduces aggregate productivity, that is, the amount of aggregate output that can be produced with a given amount of labor, as in Hsieh and Klenow (2009). To see this effect, consider the problem of allocating a given amount of labor of each type L_s across entrepreneurs in order to maximize aggregate output Y subject to the technology above. The efficient output allocations satisfy

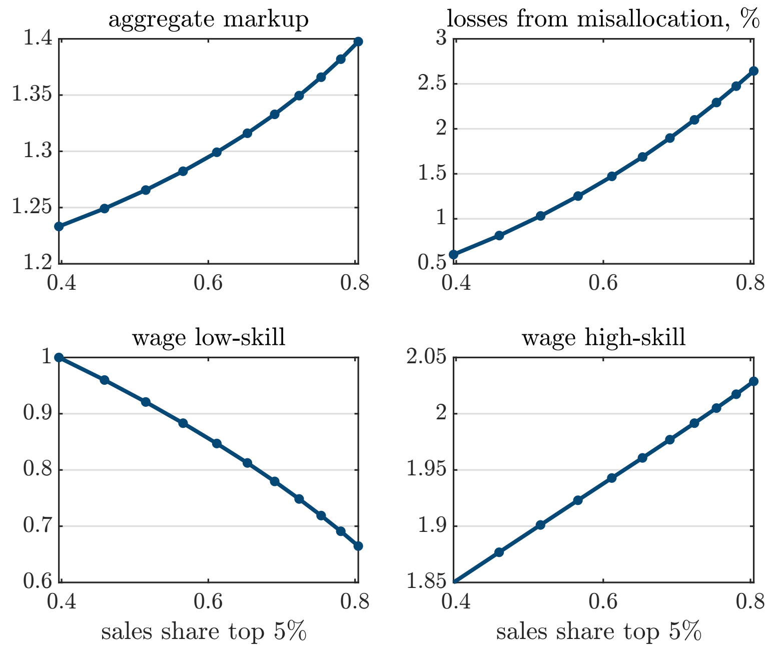
$$D\Upsilon'\left(\frac{y(z)}{Y}\right) = \frac{1}{\eta} W^*(z) \left(\frac{y(z)}{z}\right)^{\frac{1}{\eta}} \frac{1}{y(z)}, \tag{11}$$

where $W^*(z) = \left[(1-\psi(z))(W_1^*)^{1-\rho} + \psi(z)(W_2^*)^{1-\rho}\right]^{\frac{1}{1-\rho}}$ is a function of the multipliers on the two labor resource constraints, W_1^* and W_2^*. At the optimum, the marginal valuation of an additional unit of a variety is equal to the marginal cost of producing it. In contrast, as shown in equation (7), the quantity chosen by the entrepreneur equates the marginal valuation of the variety (its price) to a markup over the marginal cost. If markups vary across firms, aggregate output is below the efficient level. Our calibration implies that the losses from misallocation are equal to 0.6%.

In addition to production consequences, in our economy with heterogeneous households, markups also have distributional consequences: they redistribute income across households. To see this, notice that the income share of the two types of workers is $\frac{\mathcal{E}_s}{\mathcal{M}}$ and decreases in the aggregate markup, while the income share of entrepreneurs is $1 - \frac{\eta}{\mathcal{M}}$ and increases in the aggregate markup. Thus, markups redistribute income from workers to entrepreneurs. Moreover, since higher ability entrepreneurs earn higher markups and profits, dispersion in markups increases inequality among entrepreneurs. Finally, because high-markup firms hire disproportionately more high-skill workers, markups also affect the skill premium and the degree of inequality among workers.

To illustrate the mechanism of the model, Figure 1 traces out the impact of higher product market concentration, as measured by the sales share of the largest 5% of firms, on macro aggregates. We increase product market concentration by increasing the variance of entrepreneurial ability and adjusting the mean to keep the efficient level of aggregate productivity unchanged. The top panels illustrate the production consequences of higher concentration: it increases the aggregate markup and correspondingly lowers the labor share $\eta/\mathcal{M}$,

Figure 1: Effect of Product Market Concentration



Notes: The figure traces out the effect of increasing the variance of entrepreneurial ability on equilibrium outcomes. The x-axis reports the resulting sales share of the top 5% of firms. The y-axes report macroeconomic outcomes. We normalize the wage of low-skill workers to 1 in the baseline model.

and it increases the losses from misallocation. The bottom panels illustrate the effects on equilibrium wages: higher concentration reduces the wage of low-skill workers and increases the wage of high-skill workers, thereby increasing inequality between workers. The wage of low-skill workers falls for three reasons: *(i)* the aggregate markup increases, reducing the demand for labor, *(ii)* aggregate productivity falls, reducing the marginal product of labor, and *(iii)* firms reduce the demand for low-skill workers in favor of high-skill workers ($\mathcal{E}_1$ falls and $\mathcal{E}_2$ increases). The latter effect is strong enough to bid up the wage of high-skill workers despite the rise in markups and decline in productivity. Thus our model is consistent with the evidence from the literature on superstar firms that the reallocation of production from firms with high labor shares (low markups) to firms with low labor shares (high markups) depresses the aggregate labor share (Autor et al., 2017, Kehrig and Vincent, 2021).

2.4 Regulator's Problem

We consider the problem of a regulator who designs optimal product market interventions: it chooses how to allocate production $y(z)$ and consumption $c(z)$ across entrepreneurs, recognizing that its prescription determines aggregate output Y and the equilibrium wages W_s.

Following the Mirrleesian approach to optimal taxation, we assume that the regulator does not observe the ability of individual entrepreneurs and thus faces incentive compatibility constraints. We characterize the optimal allocations chosen by the regulator under the assumption that the interventions are *revenue-neutral* so that the net amount of transfers to entrepreneurs is equal to zero. Because the regulator can only intervene in the product market, it can affect the welfare of workers only by changing the equilibrium wages. We derive a Diamond-Saez-type formula that describes the optimal allocations and study the resulting implications for the optimal degree of product market concentration.

2.4.1 Incentive Compatibility Constraints

Because the regulator does not know an individual entrepreneur's ability, its choice of consumption $c(z)$ and output $y(z)$ must satisfy incentive compatibility constraints. Let $\tau(z)$ be a transfer received by an entrepreneur who claims to have ability z. The entrepreneur's consumption when it truthfully reveals its type is

$$c\left(z\right)=D\Upsilon'\left(\frac{y\left(z\right)}{Y}\right)y(z)-W\left(z\right)\left(\frac{y\left(z\right)}{z}\right)^{\frac{1}{\eta}}+\tau\left(z\right), \tag{12}$$

where the first two terms are the revenue net of the labor costs. If this entrepreneur instead reports ability $\hat{z}$, it receives a transfer $\tau\left(\hat{z}\right)$ and consumption

$$c\left(\hat{z},z\right)=D\Upsilon'\left(\frac{y\left(\hat{z}\right)}{Y}\right)y(\hat{z})-W\left(z\right)\left(\frac{y\left(\hat{z}\right)}{z}\right)^{\frac{1}{\eta}}+\tau\left(\hat{z}\right). \tag{13}$$

Without loss of generality we invoke the revelation principle and focus on a truth-telling mechanism, so the incentive compatibility constraints are

$$c(z,z)\geq c(\hat{z},z)\quad\text{for all } z,\hat{z}. \tag{14}$$

We pursue a first-order approach and replace the global constraints in (14) with the local constraints

$$\left.\frac{\partial c\left(\hat{z},z\right)}{\partial\hat{z}}\right|_{\hat{z}=z}=0. \tag{15}$$

We then verify numerically that the solution to this relaxed problem indeed satisfies the global constraints in equation (14).[11] The local incentive constraints imply that the entrepreneur's

[11] The global constraints are satisfied as long as $y'(z)\geq0$ and the single-crossing property $\frac{\partial^2\pi(y,z)}{\partial y\partial z}\geq0$ is satisfied. The latter holds as long as $\frac{W'(z)z}{W(z)}\leq\frac{1}{\eta}$, which is indeed satisfied in our parameterizations.

consumption varies with productivity according to

$$c'(z) = \frac{1}{\eta}\frac{W(z)}{z}\left(\frac{y(z)}{z}\right)^{\frac{1}{\eta}} - W'(z)\left(\frac{y(z)}{z}\right)^{\frac{1}{\eta}}. \tag{16}$$

The first term captures the fact that a more productive entrepreneur needs less labor to produce a given amount of output. These cost savings translate into higher consumption, more so the larger the equilibrium wages. The second term captures the fact that a more productive entrepreneur hires more high-skill labor, and therefore faces a higher composite wage $W(z)$. This effect dampens the increase in consumption and loosens the incentive compatibility constraint. Since we restrict attention to parameterizations in which the shape of $\psi(z)$ implies that $W(z)/z^{\frac{1}{\eta}}$ decreases with z (otherwise more productive entrepreneurs would produce less), $c'(z) > 0$, so more productive entrepreneurs earn information rents and enjoy more consumption.

2.4.2 Optimal Regulation

Assuming a utilitarian objective and letting

$$V_s^w(W_s) = \omega_s \int_0^\infty v_s(e, W_s)\, \mathrm{d}H(e) \tag{17}$$

denote the welfare of workers of type s, defined in equation (2), the problem of a utilitarian regulator is to choose $y(z)$, $c(z)$, Y, W_1 and W_2 to maximize

$$V_1^w(W_1) + V_2^w(W_2) + \omega \int_0^\infty \frac{c(z)^{1-\theta}}{1-\theta}\, \mathrm{d}F(z) \tag{18}$$

subject to the incentive compatibility constraints (16), the aggregate production function

$$\omega \int_0^\infty \Upsilon\left(\frac{y(z)}{Y}\right) \mathrm{d}F(z) = 1, \tag{19}$$

the labor resource constraints

$$\omega \int_0^\infty (1-\psi(z))\left(\frac{W_s}{W(z)}\right)^{-\rho}\left(\frac{y(z)}{z}\right)^{\frac{1}{\eta}} \mathrm{d}F(z) = L_s^w(W_s),\ \forall s, \tag{20}$$

and the aggregate resource constraint

$$C_1^w(W_1) + C_2^w(W_2) + \omega \int_0^\infty c(z)\, \mathrm{d}F(z) = Y, \tag{21}$$

which follows from our requirement that the regulator's interventions are revenue-neutral, so that $\int_0^\infty \tau(z)\, \mathrm{d}F(z) = 0$.[12]

[12] Note that the individual rationality constraints do not bind here because entrepreneurs have no other source of income and have preferences that satisfy the Inada conditions.

We show in the Appendix that the solution to the regulator's problem is characterized by the following condition that determines the output across producers

$$D\Upsilon'\left(\frac{y(z)}{Y}\right)=\left(1+\underbrace{\mu(z)\frac{\left[\frac{1}{\eta}\frac{W(z)}{z}-W'(z)\right](1-F(z))}{\nu(z)f(z)}}_{\xi(z)}\right)\frac{\nu(z)}{W(z)}\frac{1}{\eta}W(z)\left(\frac{y(z)}{z}\right)^{\frac{1}{\eta}}\frac{1}{y(z)}. \tag{22}$$

As in the decentralized allocation in equation (7), the regulator introduces a wedge between the price $D\Upsilon'\left(\frac{y(z)}{Y}\right)$ and the marginal cost $\frac{1}{\eta}W(z)\left(\frac{y(z)}{z}\right)^{\frac{1}{\eta}}\frac{1}{y(z)}$. This wedge is the product of two terms: one, $1+\xi(z)$, arises from the information frictions, and the other, $\nu(z)/W(z)$, from the fact that the skill premium does not coincide with the regulator's relative valuation of the two types of labor.

Consider first the term that captures the information frictions and shares many similarities to that in the Mirrleesian taxation literature.[13] The term

$$\mu(z)=1-\frac{1}{\lambda}\frac{1}{1-F(z)}\int_z^\infty c(x)^{-\theta}f(x)\,\mathrm{d}x$$

depends on the ratio of the average marginal utility of consumption of entrepreneurs with ability above z, namely $\frac{1}{1-F(z)}\int_z^\infty c(x)^{-\theta}f(x)\,\mathrm{d}x$, to the regulator's valuation of an additional unit of consumption, λ, and therefore captures the desire to redistribute. The term $\left[\frac{1}{\eta}\frac{W(z)}{z}-W'(z)\right](1-F(z))$ represents the amount of consumption that the regulator can collect from all entrepreneurs with ability greater than z by distorting the production of entrepreneurs with ability equal to z. To understand why this is the case, note that the incentive compatibility constraint can be rewritten as $c'(z)=\left[\frac{1}{\eta}\frac{W}{z}-W'(z)\right]l(z)$ so by marginally reducing employment for entrepreneurs with productivity z, the regulator is able to reduce the consumption of all entrepreneurs with productivity above z by $\frac{1}{\eta}\frac{W}{z}-W'(z)$ times the mass of such entrepreneurs, $1-F(z)$. The redistributive gains from distortions must be balanced against the output losses from reducing employment. Since the mass of producers with productivity z is equal to $f(z)$, these losses, evaluated at the effective marginal rate of substitution between composite employment at a firm with productivity z and consumption,

$$\nu(z)=\left(\nu_1W_1^{-\rho}(1-\psi(z))+\nu_2W_2^{-\rho}\psi(z)\right)W(z)^{\rho},$$

amount to $\nu(z)f(z)$. Here, ν_s is the marginal rate of substitution between employment of each type and consumption, that is, the ratio of the multipliers on the labor to goods resource constraints.

[13] Diamond (1998), Saez (2001), Golosov et al. (2016), Heathcote and Tsujiyama (2021), Sachs et al. (2020).

The second term in the expression for the wedge,

$$\frac{\nu(z)}{W(z)} = \frac{\nu_1}{W_1} \frac{(1-\psi(z)) W_1^{1-\rho} + \psi(z) \frac{\nu_2/\nu_1}{W_2/W_1} W_2^{1-\rho}}{(1-\psi(z)) W_1^{1-\rho} + \psi(z) W_2^{1-\rho}},$$

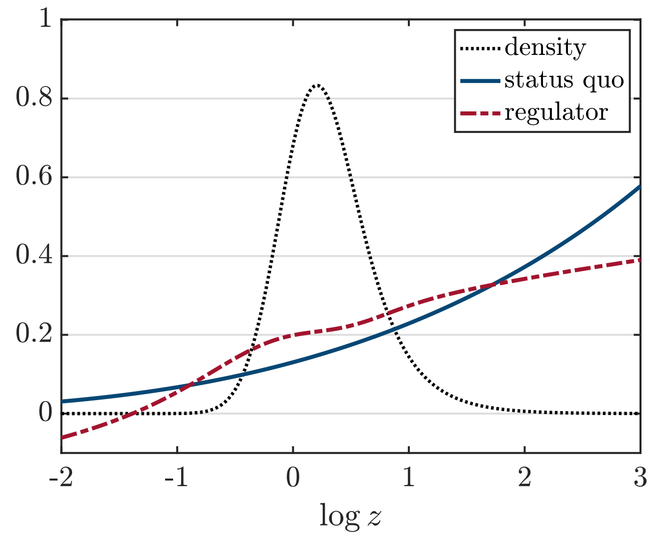
varies with z if two conditions are simultaneously satisfied: there is assortative matching between firms and workers so $\psi(z)$ varies with z and the relative shadow value of the two types of labor ν_2/ν_1 is different from the skill premium W_2/W_1. The regulator recognizes that by prescribing that more productive entrepreneurs produce more, it increases the skill premium and therefore increases inequality among workers. The regulator thus chooses to distort production relative to the efficient allocations in equation (11) in order to reduce inequality between worker types. This force is present even absent information frictions. If $\psi(z)$ were constant, the wedge $\nu(z)/W(z)$ would be constant as well, so the regulator would restore allocative efficiency absent information frictions.

Figure 2 illustrates the wedge between price and marginal cost chosen by the regulator and contrasts it with the markup, the wedge in the status quo.[14] The optimal wedge is upward-sloping, reflecting that equity considerations dominate efficiency concerns. Importantly, the optimal wedge is flatter than the markup. Since allocative efficiency requires that the wedge is constant across firms, the regulator's allocations feature less misallocation and therefore a higher level of aggregate productivity.

Table 3 summarizes the implications of optimal regulation for product market concentration as measured by the sales share of the largest producers. We also report the degree of product market concentration implied by the efficient allocations in equation (11) that maximize aggregate output and therefore entirely eliminate the wedge between price and marginal cost. By eliminating the markup wedge, the efficient allocation implies higher product market concentration than in the status quo. For example, the largest 10% of firms account for 52% of sales in the status quo and 58% under the efficient allocations. As shown in equation (22), a regulator with concerns for redistribution does not restore allocative efficiency. Nevertheless, optimal regulation also implies higher product market concentration than in the status quo. For example, under the allocations chosen by the utilitarian regulator, the largest 10% of firms account for 54% of sales.

[14] As pointed out by Heathcote and Tsujiyama (2021), the solution to Mirrleesian optimal tax problems can be highly sensitive to the number of nodes used in discretization. We therefore solve the system of differential equations that characterize the optimal allocations using 25,000 Gauss-Legendre nodes and weights to discretize the distribution of ability. Increasing the number of nodes to 100,000 makes no meaningful difference to the results.

Figure 2: Wedge Between Price and Marginal Cost



Notes: For visual clarity, we truncate the range of ability z. We super-impose the density of z, scaled by a constant.

Table 3: Product Market Concentration

	Status quo	Efficient allocation	Optimal regulation utilitarian	Optimal regulation only workers
Sales share top 1%	0.19	0.23	0.21	0.23
Sales share top 5%	0.40	0.45	0.42	0.46
Sales share top 10%	0.52	0.58	0.54	0.60

One may conjecture that the result that optimal regulation prescribes more product market concentration is driven by the fact that the utilitarian objective also incorporates the welfare of entrepreneurs. We show that this is not the case. To that end, the last column of Table 3 reports the product market concentration prescribed by a regulator that seeks to only maximize the welfare of the two types of workers and places zero weight on the welfare of entrepreneurs. Perhaps counterintuitively, product market concentration is higher in this case, even higher than is required to restore allocative efficiency: the largest 10% of firms account for 60% of sales. We explain the intuition for this result below.

2.4.3 Implementation and Intuition

We next provide intuition for our results above. We do so in two steps. First, we show that the optimal product market interventions are well-approximated by an output subsidy function characterized by three parameters that can be intuitively interpreted. Second, we provide intuition for our mechanism by tracing out the impact of increasing the slope of the marginal subsidy schedule on product market concentration and equilibrium outcomes.

Restricted Subsidy. The allocations chosen by the regulator can be implemented by introducing a production subsidy that specifies the after-subsidy revenue $S(y)$ of an entrepreneur who produces y units of output.[15] We show that a simple parametric subsidy function

$$\hat{S}(y) = \tau_0 + \frac{\tau_1}{1+\tau_2} \Upsilon \left(\frac{y}{Y}\right)^{1+\tau_2} \tag{23}$$

can achieve most of the welfare gains attainable by the unrestricted schedule $S(y)$. Here τ_0 determines the lump-sum transfer, τ_1 determines the average level of marginal subsidies and τ_2 determines the slope of the marginal subsidy schedule.

The wedge $\hat{m}(z)$ between price and marginal cost implied by this subsidy schedule is

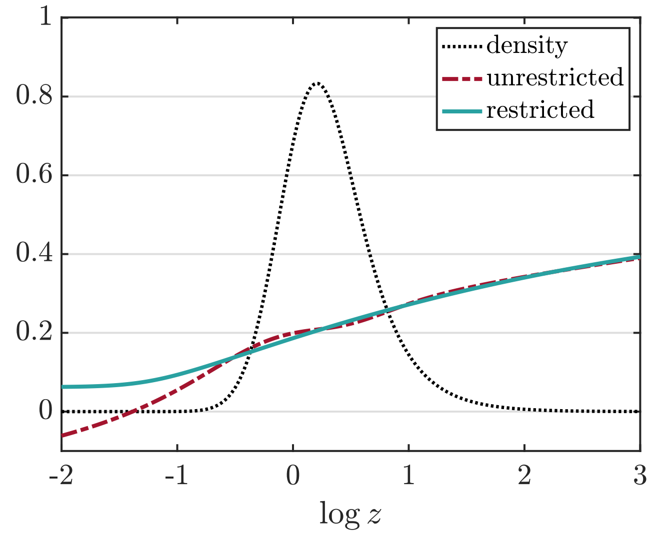
$$\hat{m}(z) = \frac{1}{\tau_1 \Upsilon \left(\frac{y(z)}{Y}\right)^{\tau_2}}.$$

If $\tau_2 = 0$, the wedge does not depend on productivity, so this subsidy function restores allocative efficiency. Moreover, since $\Upsilon(\cdot)$ is an increasing function, the wedge declines (increases) with a producer's output whenever $\tau_2 > 0$ (< 0). Thus, τ_2 determines how a given amount of labor is allocated *across* producers and the overall degree of product market concentration. In turn, τ_1 determines the aggregate wedge between price and marginal cost, and therefore the overall demand for labor. Finally, given τ_1 and τ_2, τ_0 adjusts to ensure revenue neutrality.

Figure 3 shows that the wedge between price and marginal cost implied by the unrestricted optimal policy chosen by the utilitarian regulator closely aligns with that implied by the optimally chosen restricted subsidy. The optimal level of τ_2 is -0.025, so the wedge increases with firm size, implying a lower degree of product market concentration compared to the efficient allocations. Because the wedges under the restricted and unrestricted subsidies are nearly the same, the implications for product market concentration and welfare are also nearly identical. For example, the sales share of the largest 10% of firms is 54% and the

[15] We implement the optimal allocations using a quantity subsidy $S(y)$, rather than a sales subsidy $S(p(y)y)$, because the demand elasticity of the most productive entrepreneurs may fall below one at the optimal allocation. Since in this region sales fall with the quantity produced, there does not exist a single-valued sales subsidy function that implements the regulator's optimal allocations. Additionally, the assumption that entrepreneurial ability is private information precludes conditioning $S(\cdot)$ on profits.

Figure 3: Wedges Under Restricted and Unrestricted Subsidy Schedule



consumption equivalent welfare gains from implementing the optimal regulation are 1.1% under both the restricted and unrestricted schedules.[16]

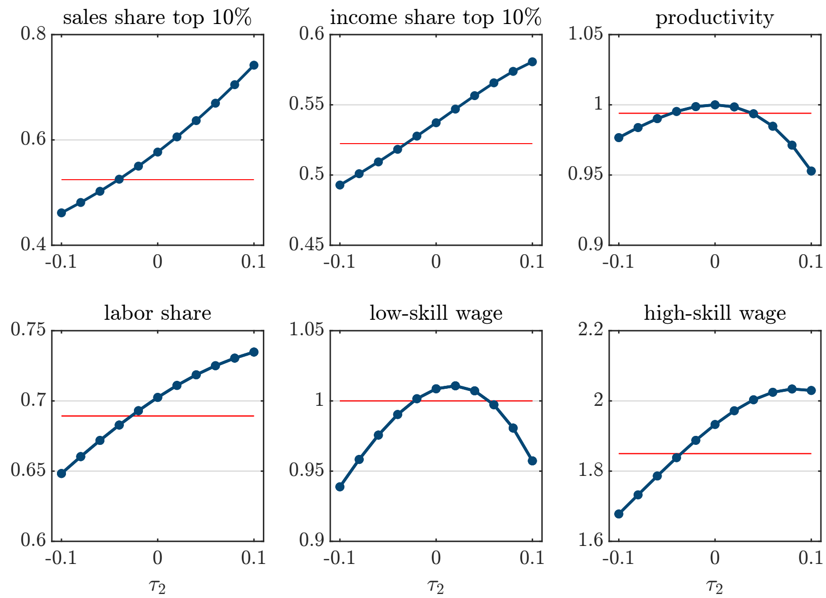
Intuition. We find this simple three-parameter subsidy schedule useful because it allows us to provide sharper intuition for the tradeoffs the regulator faces in deciding how to intervene in the product market. Consider the following experiment in which, for clarity, we fix the lump-sum transfer at zero and trace out the implications of increasing τ_2 while reducing τ_1 to ensure revenue neutrality.

As the top panels of Figure 4 show, a higher τ_2, which implicitly subsidizes larger producers at the expense of smaller ones, increases product market concentration and leads to higher income inequality. Income inequality increases both because productive entrepreneurs are subsidized and earn higher profits, as well as because the skill premium increases. Changing τ_2 also affects aggregate productivity: since $\tau_2 = 0$ recovers the efficient allocations, aggregate productivity is maximized at this point.

The bottom panels of the figure trace out the implications of varying τ_2 for labor market outcomes. Increasing τ_2 bids up the demand for labor and increases the labor share. To see why that is the case, we note that in the presence of production subsidies the aggregate labor

[16]We calculate the consumption-equivalent welfare gains using the approach of Benabou (2002). Specifically, we first calculate the constant amount of consumption $\bar{c}$ every household would have to receive so that society achieves the same level of utilitarian welfare as under the equilibrium allocations. We then define the welfare gains as the percent change in $\bar{c}$. See Boar and Midrigan (2022a) for details.

Figure 4: Effect of Varying τ_2



Notes: The horizontal lines represent the values of the variables in the status quo.

share can be written as

$$\frac{W_1 L_1 + W_2 L_2}{Y} = \eta \frac{\mathcal{S}}{\mathcal{M}},$$

where $\mathcal{S}$ is a weighted average of the producer-level subsidies and $\mathcal{M}$ is the aggregate markup, now computed as a ratio of price to marginal cost *inclusive* of the subsidy. Even though the aggregate markup increases due to the increase in the market share of the largest, high markup firms, the subsidy more than offsets the increase in markup, thus reducing the wedge in the firms' optimality conditions for labor and raising the labor share.

The following example may further clarify the intuition for why subsidies that encourage producers to expand lead to an increase in the labor share. Consider an economy with a unit mass of identical firms with technology $y = l^\eta$ and an intervention that changes the after-subsidy revenue to $\frac{\tau_1}{1+\tau_2} y^{1+\tau_2}$. Suppose that τ_1 is set to ensure revenue neutrality, so that the post-subsidy aggregate revenue, $\frac{\tau_1}{1+\tau_2} Y^{1+\tau_2}$, is equal to aggregate output, Y. Such an intervention changes the labor share to $WL/Y = \eta(1+\tau_2)$. If $\tau_2 > 0$, this size-dependent subsidy effectively increases the span-of-control to $\eta(1+\tau_2)$, reducing the income share of producers and increasing the labor share.

A higher τ_2 does not uniformly increase the demand for low- and high-skill labor in our economy with assortative matching. Because more productive firms hire disproportionately more high-skill workers, the skill premium increases. Nevertheless, the wages of low-skill workers are maximized at a value of τ_2 that is greater than zero, which explains why a regulator that only values the welfare of workers prescribes a value of τ_2 larger than that required to restore allocative efficiency.

To summarize, product market interventions that require larger producers to expand have both benefits and costs. On the cost side, they lead to more product market concentration, which increases markups, and to more inequality both within and across groups. On the benefits side, starting from an economy that is distorted to begin with, they increase productivity by improving allocative efficiency and they bid up the demand for labor, thus increasing the labor share. The larger the regulator's weight on the welfare of workers, the more the benefits outweigh the costs, and therefore the larger the degree of product market concentration chosen by the regulator.

2.5 Robustness

In the Appendix we explore how our result regarding the implications of optimal regulation for product market concentration varies with key model parameters. We find that optimal regulation leads to more product market concentration the higher the super-elasticity ε/σ and the lower the demand elasticity σ, that is, the higher the markup distortions are. When markup distortions are small enough, a utilitarian regulator may prescribe lower product market concentration, in an effort to redistribute to poor entrepreneurs. A regulator who only seeks to maximize the welfare of workers always increases product market concentration, typically above levels required to restore allocative efficiency. We also find that optimal regulation leads to more product market concentration the lower is the share ω of entrepreneurs in the population. For very high values of ω, above empirically plausible ones, a utilitarian regulator may prescribe lower product market concentration because the motive to redistribute to poor entrepreneurs becomes stronger. A regulator who only values the welfare of workers does not have such a motive and unambiguously increases product market concentration. The conclusion that optimal regulation increases product market concentration is also robust to changing the variance of labor market and entrepreneurial ability.

In the Appendix we also study an alternative economy in which firms have labor market power. They thus pay their workers a wage that is a markdown over their marginal product of labor, with this markdown decreasing with firm size. In this economy, as in the economy with markups, more concentration decreases the labor share. Nevertheless, optimal regulation once again encourages larger producers to expand and leads to more product market concentration.

3 Dynamic Model

We have purposely abstracted above from a number of features in order to highlight the key tradeoffs between equity and efficiency entailed by product market interventions. We next enrich the model by introducing three additional ingredients. First, we allow for capital and wealth accumulation to study the implications of product market interventions for wealth inequality. Second, we assume that entrepreneurs co-exist with corporate firms so that the ownership of firms is more diversified compared to our static model. Third, we assume a government that provides some redistribution via taxes and transfers. We use this setting to study the optimal degree of product market interventions in the restricted class considered above and show that our main result that optimal regulation features a greater degree of product market concentration than under the status quo is robust in this richer setting. We abstract from aggregate uncertainty and study optimal unanticipated policy reforms, taking into account the transition dynamics between steady states.

3.1 Households

Workers seek to maximize their life-time utility given by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{c_t^{1-\theta}}{1-\theta} - \frac{h_t^{1+\gamma}}{1+\gamma} \right)$$

subject to the budget constraint

$$c_t + a_{t+1} = i_t - T\left(i_t\right) + a_t,$$

where a_{t+1} are savings, $T\left(\cdot\right)$ is the income tax schedule and

$$i_t = r_{t-1} a_t + W_{st} e_t h_t$$

is pre-tax income, derived from the return on asset holdings and from working.

Entrepreneurs maximize life-time utility

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\theta}}{1-\theta}$$

subject to an identical budget constraint as that of the workers. Their income

$$i_t = r_{t-1} a_t + \pi_t\left(z_t\right)$$

derives from the return on wealth and profits from the business.

All households save with perfectly competitive financial intermediaries at a risk-free rate r_t. Financial intermediaries use the resources obtained from households to purchase capital, shares in corporate firms and a risk-free government bond. The income tax schedule is

$$T(i_t) = i_t - (1-\tau)\frac{i_t^{1-\xi}}{1-\xi} - \iota_t, \tag{24}$$

where τ governs the level and ξ the slope of the marginal tax schedule, while ι_t is a lump-sum transfer. This specification has been shown to approximate well the U.S. tax and transfer system (Heathcote et al., 2017 and Boar and Midrigan, 2022a). We assume that labor efficiency and entrepreneurial ability follow independent Markov processes.

3.2 Final Good Firms

The final good Y_t is used for consumption, C_t, investment in physical capital and the creation of corporate firms X_t, and government spending G, so the aggregate resource constraint is

$$Y_t = C_t + X_t + G. \tag{25}$$

As earlier, the final good is assembled using the Kimball (1995) production function

$$\int_0^{N_F} \Upsilon\left(\frac{y_{it}}{Y_t}\right) \mathrm{d}i = 1,$$

where $N_F = \omega + N$ is the total mass of firms in the economy, which includes the mass ω of entrepreneurs and the mass N of corporate firms. As we describe below, the assumptions we make on the cost of creating new firms ensure that the mass of corporate firms is constant. The optimal input choices of the final good producers give rise to identical demand curves as in our static model. We implicitly assume that private businesses compete alongside corporate firms in the product market.[17]

3.3 Intermediate Goods Producers

Each variety is produced by a single producer, either corporate or privately owned. The technology with which a producer with ability z_t operates is now augmented to include capital and is

$$y_t = z_t \left(k_t^{\alpha} l_t^{1-\alpha}\right)^{\eta},$$

[17] See Smith et al. (2019), who show that the two types of firms coexist across U.S. industries. An earlier draft of our paper showed that the impact of product market interventions is similar in economies without either corporations or entrepreneurs.

where, as in the static model, l_t is a labor composite

$$l_t = \left[(1-\psi(z_t))^{\frac{1}{\rho}} l_{1t}^{\frac{\rho-1}{\rho}} + \psi(z_t)^{\frac{1}{\rho}} l_{2t}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}},$$

with $\psi(z_t)$, described in equation (4), determining the degree of assortative matching. The firm maximizes profits,

$$\pi_t = p_t(y_t)y_t - W_{1t}l_{1t} - W_{2t}l_{2t} - R_t k_t,$$

where $R_t = r_{t-1} + \delta$ is the rental rate of capital, pinned down by a no-arbitrage condition. As in the static model, the firm sets a price equal to a markup over its marginal cost and the markup is increasing in its market share.

Corporate and privately-held firms produce with identical technology, so they only differ in their ownership structure and tax treatment. Unlike private firms, which are pass-through businesses, corporate firms are subject to a corporate profit tax. For ease of exposition only, we assume that the productivity of corporate firms is constant over time.[18]

Corporate firms exit with exogenous probability φ, so their mass evolves according to

$$N_{t+1} = (1-\varphi)(N_t + \vartheta_t),$$

where ϑ_t is the mass of entrants. The free entry condition requires that the cost $\mathcal{K}_t$ of creating a new firm, denominated in units of the final good, is equal to the expected value Q_t of a new firm,

$$\mathcal{K}_t = Q_t \equiv \int Q_t(z)\, \mathrm{d}F^c(z),$$

where

$$Q_t(z) = \frac{1-\varphi}{1+r_t} \left[Q_{t+1}(z) + (1-\tau_c)\pi_{t+1}(z) \right]$$

is the price of a claim to the after-tax profits of a firm with productivity z, and τ_c is the corporate profit tax rate. Upon entering, a corporate firm draws its productivity from a distribution $F^c(z)$, so the expected return to entry is equal to $\int Q_t(z)\, \mathrm{d}F^c(z)$.

We follow Gutiérrez et al. (2021) in assuming that entry costs increase with the mass of entrants. Specifically,

$$\mathcal{K}_t = \left(\bar{\mathcal{K}}\vartheta_t\right)^{\frac{1}{\phi}},$$

where $\bar{\mathcal{K}}$ determines the average level of entry costs, and ϕ determines the elasticity of entry rates to changes in the value of corporate firms. In our quantitative analysis we assume

[18] This assumption is without loss of generality since the ownership of these firms is fully diversified and only the stationary distribution of their productivity matters for equilibrium outcomes.

$\phi \to 0$, so that in equilibrium the mass of entrants is constant.[19] This is a conservative assumption because it implies the largest response of stock prices to policies that increase product market concentration and thus an upper bound on the distributional costs that determine the equity-efficiency tradeoff of such policies.[20]

3.4 Government and Financial Intermediaries

The government issues a time-invariant stock of debt B. It finances interest on this debt and an exogenously given amount of government spending G using personal income and corporate profit taxes, T_t^i and T_t^c, so its budget constraint is

$$r_{t-1}B + G = T_t^i + T_t^c.$$

For notational convenience, we assume that households deposit their savings with financial intermediaries who use these resources to purchase capital, government bonds and shares in corporate firms. Absent aggregate uncertainty, the rate of return on all these assets is identical and denoted by r_t.

3.5 Equilibrium

An equilibrium consists of *(i)* aggregate prices W_{st}, R_t, r_t, Q_t, *(ii)* consumption, saving and labor supply choices of workers $c_{st}(a,e)$, $a_{st+1}(a,e)$, $h_{st}(a,e)$, *(iii)* consumption and savings choices of entrepreneurs $c_t(a,z)$, $a_{t+1}(a,z)$, *(iv)* employment, capital, output and price choices of producers $l_{st}(z)$, $k_t(z)$, $y_t(z)$, $p_t(z)$, *(v)* measures of workers $n_{st}(a,e)$ and entrepreneurs $n_t(a,z)$ over their idiosyncratic states, *(vi)* mass of corporate firms N_t, such that

1. Given prices, the households' decisions maximize their life-time utility and the production choices maximize firm profits.

2. Aggregate output satisfies the Kimball aggregator

$$\int \Upsilon\left(\frac{y_t(z)}{Y_t}\right) \mathrm{d}n_t(a,z) + N_t \int \Upsilon\left(\frac{y_t(z)}{Y_t}\right) \mathrm{d}F^c(z) = 1.$$

3. Markets clear period by period. The labor market clearing conditions are

$$\int l_{st}(z)\ \mathrm{d}n_t(a,z) + N_t \int l_{st}(z)\ \mathrm{d}F^c(z) = \int e h_{st}(a,e)\ \mathrm{d}n_{st}(a,e),\ s = 1,2.$$

[19] Note that as $\phi \to 0$, $\mathcal{K}_t \to 0$ if $\vartheta_t < \frac{1}{\mathcal{K}}$, and $\mathcal{K}_t \to \infty$ if $\vartheta_t > \frac{1}{\mathcal{K}}$, so $\vartheta_t = \frac{1}{\mathcal{K}}$ in equilibrium.

[20] See our earlier draft, Boar and Midrigan (2019), which considers the opposite scenario with perfectly elastic entry and no changes in stock prices, and obtains similar implications of product market interventions.

The asset market clearing condition is

$$A_{t+1} = K_{t+1} + Q_t \left(N_t + \vartheta_t\right) + B,$$

where

$$A_{t+1} = \sum_{s=1}^{2} \int a_{st+1}\left(a,e\right) \mathrm{d}n_{st}\left(a,e\right) + \int a_{t+1}\left(a,z\right) \mathrm{d}n_t\left(a,z\right)$$

is the aggregate supply of assets and

$$K_t = \int k_t\left(z\right)\, \mathrm{d}n_t\left(a,z\right) + N_t \int k_t\left(z\right)\, \mathrm{d}F^c\left(z\right)$$

is the aggregate capital stock.

The goods market clearing condition in (25) is satisfied by Walras' Law. We note that investment X_t includes both investment in physical capital, $K_{t+1} - (1-\delta)K_t$ and the cost of creating new corporate firms $\mathcal{K}_t\vartheta_t$.

4. The budget constraint of the government is satisfied period by period.

5. The law of motion for the measures $n_{st}\left(a,e\right)$ and $n_t(a,z)$ evolve according to an equilibrium mapping dictated by the households' optimal savings choices and the stochastic processes for labor market efficiency and entrepreneurial ability.

6. The mass of corporate firms evolves according to $N_{t+1} = (1-\varphi)\left(N_t + \vartheta_t\right)$ and the free entry condition is satisfied.

3.6 Parameterization

We next describe how we choose parameters for our quantitative analysis. We assume the economy is in a steady state in 2013, so we target statistics for this year.

Assigned Parameters. We assign the same values to the parameters that are common across the dynamic and the static model, listed in the left column of Table 1. The new parameters are the elasticity of capital in production α, which we set to 1/3, the depreciation rate δ, which we set to 0.06, the exit rate of corporate firms φ, which we set equal to 0.04 to match that exiting firms account for approximately 4% of employment according to the Statistics of US Businesses.[21] We list these assigned parameters in the left panel of Table 4.

[21] Since we abstract from aggregate uncertainty and therefore equity premia, absent exit and therefore entry the model would greatly overstate the market value of corporate firms.

To set the tax parameters, we use the estimates of the income tax function from Boar and Midrigan (2022a), $\tau = 0.255$, $\xi = 0.049$ and $\iota = 0.164$. These are derived from the CBO data on pre- and post-tax income and imply that the median marginal tax rate is 0.27, the marginal tax rate at the 95^{th} percentile is 0.34, and the lump-sum transfer is 0.16 of GDP. We set the corporate profit tax τ_c to 0.36, consistent with the United States tax code (Bhandari and McGrattan, 2020). The unanticipated product market interventions we consider give rise to one-time unexpected capital gains due to the change in the rental rate of capital and the value of corporate firms. We assume that these are taxed at a rate of $\tau_k = 0.20$, the capital gains tax in the United States in 2013. Finally, we set the stock of government debt B equal to 100% of GDP, as in the US data.

Table 4: Parameter Values in Dynamic Model

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	0.127	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	0.110	std. dev. z shocks
ρ	1.41	skill elasticity of substitution	σ	8.717	demand elasticity at $q = 1$
ω	0.12	fraction of entrepreneurs	μ_c	1.754	mean productivity corporations
ω_2	0.28	fraction high-skill workers	$\bar{\mathcal{K}}$	0.076	entry costs / GDP
ε/σ	0.15	super-elasticity of demand	$\bar{\psi}$	2.945	avg. elasticity of high-skill labor
φ	0.04	exit rate, corporations	ζ	0.635	sensitivity of elasticity high-skill labor

Calibrated Parameters. We choose the remaining parameters, reported in the right panel of Table 4, to match salient facts about wealth and income inequality, assortative matching and the relative size of the corporate sector, reported in Table 5. We assume that labor market and entrepreneurial ability follow independent AR(1) processes with persistence ρ_e and ρ_z and Gaussian innovations with standard deviation σ_e and σ_z, respectively. As in the static model, we assume the same process for labor market ability for the two types of workers. We assume that the productivity of corporate firms is drawn from a normal distribution with mean μ_c and variance $\sigma_z^2/(1 - \rho_z^2)$, the unconditional variance of entrepreneurial productivity. The parameter μ_c thus determines how much more productive and larger are corporate firms.[22]

We choose μ_c and the cost of creating a new corporate firm $\bar{\mathcal{K}}$ to match the 63% sales share of C-corporations and the 5% share of businesses that are C-corporations in the data, as reported by Dyrda and Pugsley (2018) for 2012, the latest year in their sample. In total, investment in creating new firms amounts to 7.6% of GDP. We employ the same strategy as in the static model to calibrate the parameters that determine the skill premium and the extent of assortative matching. We also target the average wealth-to-income ratio, the share of wealth and income held by entrepreneurs, and the wealth and income Gini coefficients for all households, as well as separately for entrepreneurs and workers. As Table 5 reports, the model matches the targeted moments well.

Table 5: Moments Used to Calibrate Dynamic Model

	Data	Model		Data	Model
Wealth to income ratio	6.57	6.60	Gini wealth, workers	0.83	0.82
Wealth share of entrepr.	0.46	0.47	Gini income, workers	0.59	0.59
Income share of entrepr.	0.31	0.29	Skill premium	1.85	1.85
Gini wealth, all hhs	0.85	0.84	Wage-employment elasticity, $\times 100$	1.90	1.90
Gini income, all hhs	0.64	0.65	Fraction of corporate firms	0.05	0.05
Gini wealth, entrepr.	0.78	0.72	Sales share corporate firms	0.63	0.60
Gini income, entrepr.	0.68	0.72			

3.7 Optimal Product Market Interventions

Recall that the restricted subsidy schedule considered in Section 2.4.3 captures the vast majority of the welfare gains achievable by the Mirrleesian regulator in a static setting. Motivated by this result, we consider a regulator who contemplates a once-and-for-all unanticipated product market intervention of the form in equation (23). The regulator takes into account that its intervention alters the paths for the equilibrium wages and interest rates as the economy transitions to the new steady state.

Letting $\boldsymbol{\pi} = (\tau_0, \tau_1, \tau_2)$ denote the parameters describing the intervention and $V_0(\boldsymbol{\pi})$ denote the utilitarian objective that takes into account the implied equilibrium paths for consumption and hours worked, the problem of the regulator is to choose $\boldsymbol{\pi}$ to maximize

[22] We associate entrepreneurial firms in our model with privately-held pass-through businesses in the data and corporate firms with C-corporations in the data, regardless of whether they are privately held or publicly listed. Though imperfect, this mapping allows us to capture two key distinctions between pass-through businesses and C-corporations in the data: their tax status (pass-through vs. double taxation) and the concentration of ownership. See Dyrda and Pugsley (2018) for a detailed discussion.

$V_0(\boldsymbol{\pi})$, subject to the constraint that the intervention is revenue-neutral at all dates so that subsidies on some firms are financed by taxes on other firms. We implicitly assume that the subsidy an individual firm receives does not depend on its incorporation status.

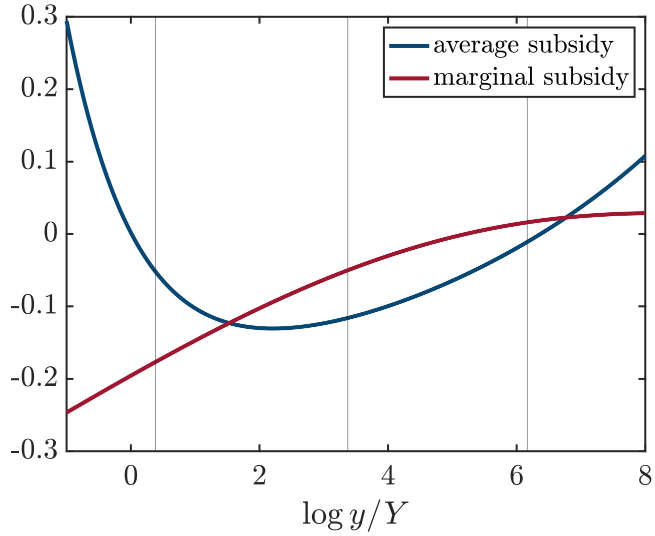
Because a particular policy intervention $\boldsymbol{\pi}$ changes equilibrium prices, it also changes the amount of revenue that the government collects in taxes. We assume that this additional revenue is rebated lump-sum to all households, so ι_t adjusts at each date to ensure that the government budget constraint is satisfied.

The policy interventions also lead to an unexpected change in the stock market value of corporate firms and in the return to capital. Even though in the steady state the rate of return to different assets is equal and individual portfolio allocations are not pinned down and are inconsequential, portfolio allocations do matter along the transition path due to the unexpected change in the stock market value of corporate firms. To ensure that these capital gains accrue to households in different parts of the wealth distribution in a way that is consistent with the portfolio holdings in the data, we use the 2013 SCF to estimate the elasticity of stock holdings to wealth. We find that this elasticity is equal to 1.08, suggesting that the wealthy hold disproportionately more stocks, consistent with Kuhn et al. (2020). We therefore assign stock holdings in the model to reproduce this empirical relationship.

Findings. A utilitarian regulator finds it optimal to set $\tau_1 = 0.716$ and $\tau_2 = 0.013$. The lump-sum transfers to producers implied by revenue neutrality amount to 2.9% of GDP. Since the values of τ_1 and τ_2 are not directly interpretable, Figure 5 reports the average and marginal subsidy schedule as a function of relative output. The average subsidy is positive at the bottom of the output distribution, reflecting the lump-sum transfer. Producers in the middle of the distribution are taxed, on average, while those at the top are subsidized, as in the static model. Optimal regulation implies that the median marginal subsidy is equal to -26.7%, while the 99^{th} percentile is equal to 1.6%. The value of τ_2 chosen by the regulator is slightly greater than zero, so it is optimal to increase product market concentration slightly more than required to implement the efficient allocations.

Table 6 reports the effect of implementing the optimal product market intervention. With the exception of the welfare gains, which take into account transition dynamics, all other statistics reflect steady-state comparisons. We start by discussing the effects of the policy chosen by the utilitarian regulator, reported in the second column of the table. Consistent with the predictions of the static model, the regulator increases product market concentration. For example, the sales share of the largest 1% of producers increases from 35% to 43%.

Figure 5: Optimal Subsidy Function: Utilitarian Regulator



Notes: The vertical bars represent the 75^{th}, 95^{th} and 99^{th} percentiles of the distribution of relative output. The average subsidy is computed as the ratio between the subsidy and sales.

Since corporate firms are larger than privately-held firms, the subsidies on large producers increase the corporate sales share from 61% to 68%. Because the regulator eliminates most misallocation, aggregate productivity increases by 1.11%.

Interestingly, optimal regulation reduces long-run inequality despite increasing concentration and the skill premium: the wealth and income Gini coefficients fall from 0.84 to 0.81 and from 0.65 to 0.62, respectively. This result reflects the loss of market share of the relatively smaller businesses owned by entrepreneurs, whose wealth and income shares fall from 0.47 to 0.32 and from 0.29 to 0.23, respectively.

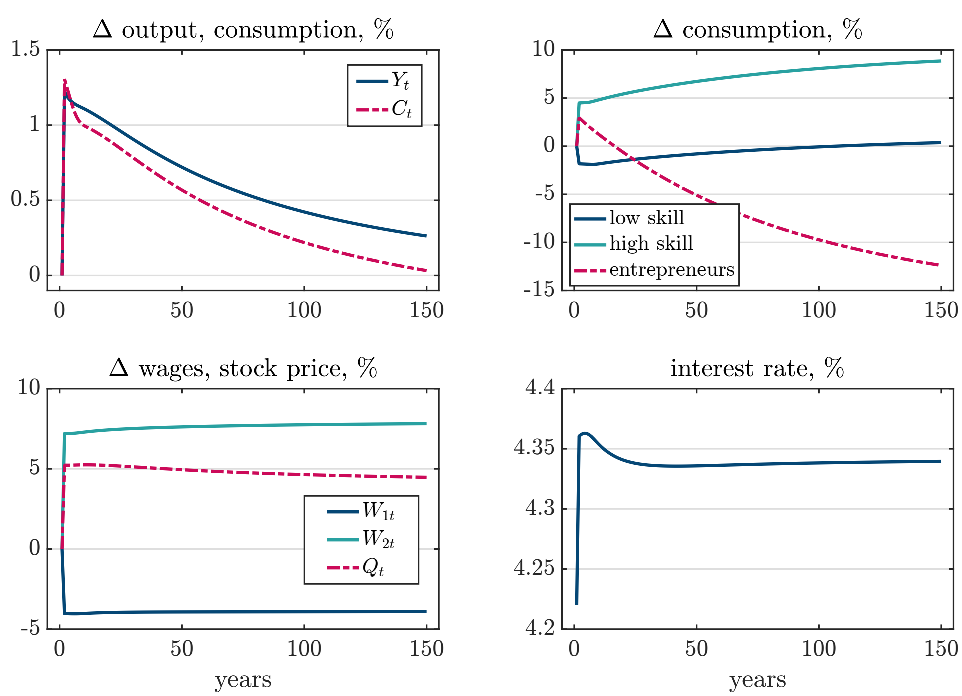
Figure 6 illustrates the transition dynamics of the key macroeconomic aggregates. Output and consumption increase on impact by 1.2% and gradually return to their pre-reform levels. This eventual decline is due to the decline in the capital stock resulting from the increase in the interest rate. Since the value of corporate firms increases after the reform, a higher interest rate is required to clear the asset market. By increasing product market concentration, optimal regulation bids up the demand for labor. The effect on wages is uneven, however, because more product market concentration increases the skill premium: the wage of high-skill workers by increases by 7.9% in the long run while that of low-skill workers falls by 3.9%. Consequently, the consumption of high-skill workers increases. While the consumption of low-skill workers falls on impact, it does not fall as much as their wages do, and ultimately recovers because the reform increases government revenue and therefore the lump-sum transfers ι_t.

Table 6: Product Market Interventions in Dynamic Model

	Status quo	Optimal regulation utilitarian	Optimal regulation only workers	Reduce concentration
Sales share top 1% firms	0.35	0.43	0.44	0.29
Sales share top 5% firms	0.79	0.85	0.87	0.68
Sales share corporations	0.60	0.68	0.70	0.51
Losses from misallocation, %	1.15	0.04	0.24	7.65
Wealth share entrepreneurs	0.47	0.32	0.35	0.48
Income share entrepreneurs	0.29	0.23	0.20	0.39
Gini wealth	0.84	0.81	0.81	0.84
Gini income	0.65	0.62	0.63	0.66
Change in welfare, %	–	3.52	2.57	−11.5
low-skill workers	–	−0.27	2.53	−15.4
high-skill workers	–	5.10	9.48	−21.2
entrepreneurs	–	21.0	−12.1	47.7
Fraction households better off	–	0.64	0.91	0.11

The consumption of entrepreneurs increases initially due to the increase in the value of their stock holdings, but eventually falls by more than 10%.

Figure 6: Transition Dynamics: Maximize Utilitarian Welfare



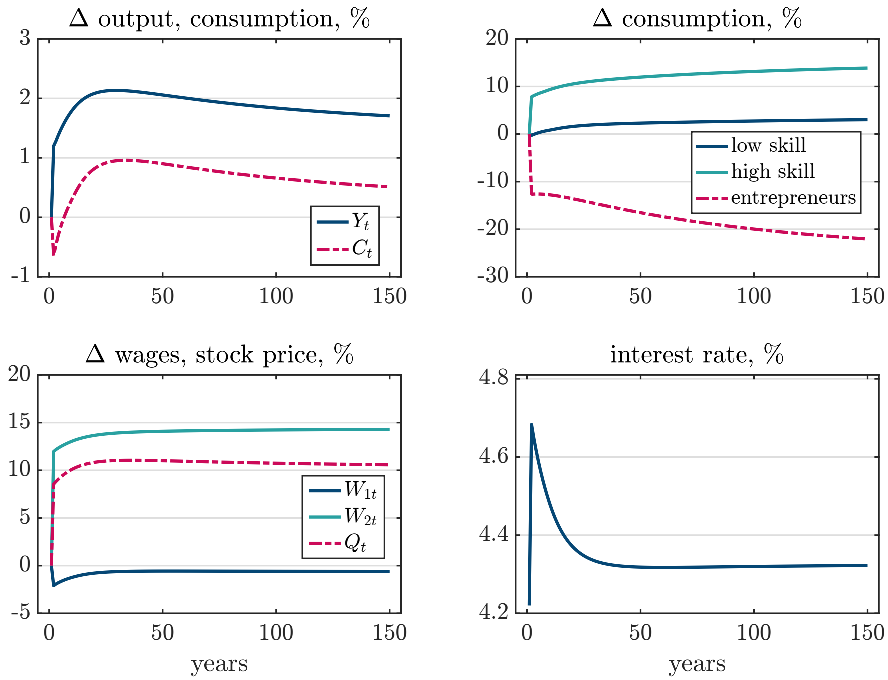
The bottom rows of Table 6 report the welfare implications of optimal regulation. Utilitarian welfare, expressed in consumption-equivalent units, increases by 3.5%.[23] Approximately two-thirds of households are better off. The gains from the reform disproportionately accrue to entrepreneurs, whose welfare increases by 21%, owing to the increase in lump-sum transfers that benefit unproductive private business owners. High-skill workers experience a 5.1% increase in welfare, whereas low-skill workers experience a modest decline in welfare of 0.3%.

As in the static model, we find it useful to contrast the optimal policy chosen by a utilitarian regulator to that chosen by a regulator who only seeks to maximize the welfare of workers and places zero weight on the welfare of entrepreneurs. We once again find that such a regulator further increases product market concentration by setting $\tau_2 = 0.024$, and thus subsidizing larger producers even more. The third column of Table 6 and Figure 7 report the effects of implementing this policy. Product market concentration increases more than under utilitarian regulation: for example, the sales share of the largest 5% of firms increases by two additional percentage points. Because such a regulator eliminates the lump-sum transfers to entrepreneurs, it is able to further subsidize production and increase the demand for labor. Consequently, the high-skill wage increases by 14% in the long run, while the low-skill wage falls by less than 1% in the long run. The majority (91%) of households benefit from this policy, with the largest gains accruing to high-skill workers (9.5% increase in welfare). Low-skill workers benefit as well (2.5% increase in welfare) due to the larger increase in output and lump-sum transfers which increase their consumption despite the drop in wages. Aggregate welfare increases by 2.6%. Not surprisingly, entrepreneurs are now worse off (12.1% drop in welfare) since the majority of them are taxed and experience a decrease in profits.

Policies that Reduce Product Market Concentration. We have shown that optimal product market interventions lead to more product market concentration, even more so when the regulator solely seeks to maximize the welfare of workers. This is because such policies bid up the demand for labor and thus increase the consumption of workers. Our results thus caution against the widely-held view that policies that reduce concentration and the market power of large firms necessarily improve the welfare of workers. Though reducing concentration indeed reduces firm market power in our model, the size-dependent interventions required to reduce the market share of large firms have the unintended consequence of

[23] As in the static model, we calculate consumption equivalent measures of welfare using the approach of Benabou (2000). We note that when preferences are logarithmic in consumption, as we assume here, these welfare gains are equivalent to the equiproportionate increase in consumption necessary to attain the same level of aggregate welfare as under the policy.

Figure 7: Transition Dynamics: Maximize Worker Welfare

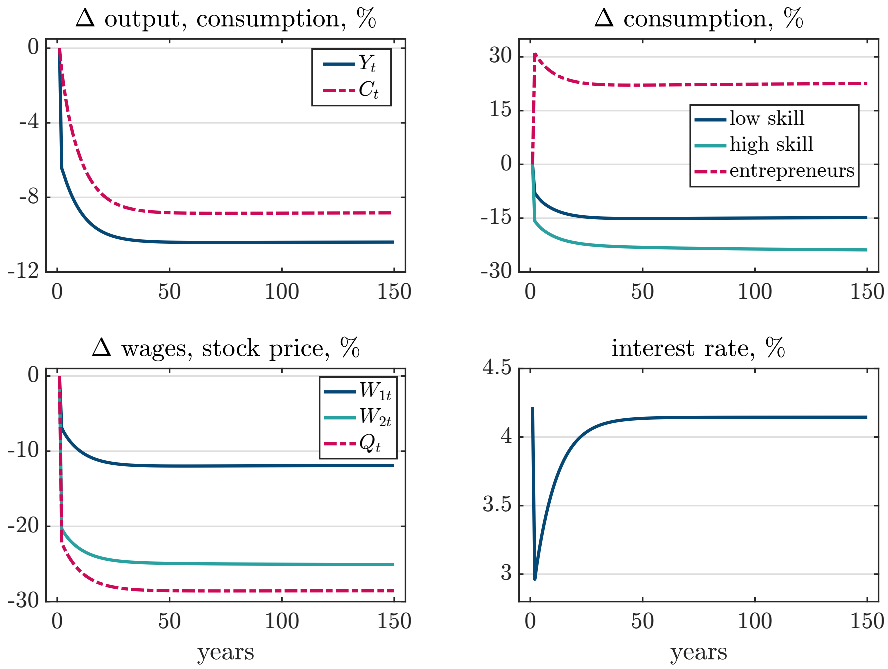


reducing the labor share, productivity and equilibrium wages.

To illustrate this point, we set $\tau_2 = -0.15$, a value that reduces the market share of the largest 5% of firms from 0.79 in the status quo to 0.68. We also eliminate the lump-sum transfer τ_0 and adjust τ_1 to ensure revenue-neutrality. The last column of Table 6 and Figure 8 show the consequences of this intervention. Output, consumption and wages, especially those of high-skill workers fall considerably, owing to both a sharp increase in misallocation (7.7% compared to 0.6% in the status quo) and because of depressed labor demand. The policy increases the income share of entrepreneurs from 0.29 to 0.39, owing to the subsidies received by heretofore smaller businesses, and leads to an increase in consumption inequality: the consumption of entrepreneurs increases by more than 20%, while that of low-skill workers falls by 15% and that of high-skill workers falls by 24%.

These results reinforce our earlier conclusions that product market concentration is not necessarily costly, even in an environment with highly unequal firm ownership. What is costly is dispersion in the marginal product of factors of production across firms and wedges between factor prices and their marginal products. Optimal regulation reduces these wedges and in doing so actually increases product market concentration. Though the largest producers benefit from such interventions at the expense of medium-sized firms, workers are better off due to higher wages.

Figure 8: Transition Dynamics: Reduce Product Market Concentration



3.8 Robustness

Our result that policies that encourage firms to expand are welfare-improving is robust to many perturbations of the model. In earlier drafts of this paper, Boar and Midrigan (2019) and Boar and Midrigan (2022b), we assumed that entrepreneurs are subject to financial constraints, that firm ownership is either perfectly diversified or fully concentrated and considered alternative ways of modelling entry into the corporate sector. We also studied a version of the model with an endogenous margin of entry into entrepreneurship, targeted an alternative set of statistics from DeBacker et al. (2023) who use IRS data on labor and business income, allowed for oligopolistic, rather than monopolistic competition across producers, as well as heterogeneity across firms in idiosyncratic distortions which break the link between firm productivity and size. Here we summarize the results of a number of additional robustness checks. For all these, unless otherwise noted, we re-calibrate the model and report the parameter values and the targeted moments in the Appendix.

We first increase the super-elasticity parameter ε/σ to 0.3, double the value in the benchmark calibration, and at the upper bound of the estimates in Edmond et al. (2023). A larger super-elasticity implies that markups increase faster with firm market shares, so the decentralized economy is more distorted. As Table 7 shows, both the utilitarian regulator, as well as the regulator who only seeks to maximize the welfare of workers, set τ_2 to a value slightly greater than zero, once again increasing product market concentration beyond what is necessary to restore allocative efficiency and above the status quo. Because this economy is more distorted by markups, the welfare gains from optimal policy are now larger. For example, the utilitarian welfare increases by 8.1% and, in contrast to the benchmark calibration, even low-skill workers are better off: their welfare increases by 3.9%.

Second, we increase the fraction of entrepreneurs ω to 20%, leaving all the other parameters unchanged. As Table 7 shows, the utilitarian regulator finds it optimal to set $\tau_2 = 0.005$, once again increasing product market concentration beyond what is required to restore allocative efficiency. Because there are more entrepreneurs now, both low- and high-skill workers experience welfare losses. A regulator who only seeks to maximize the welfare of workers raises τ_2 further, to 0.028, and is able to increase the welfare of both low- and high-skill workers by 1.9 and 9.2%, respectively.

Third, in the last two columns of Table 7 we target a wage-employment elasticity of 3.8%, twice as large as in the benchmark calibration. The utilitarian regulator chooses a $\tau_2 = 0.008$, slightly lower than in the benchmark calibration, but nevertheless subsidizes larger firms more than is required to restore allocative efficiency, increasing product market concentration. Because this calibration features more assortative matching, high-skill workers benefit disproportionately more from optimal regulation compared to low-skill workers. Nevertheless, the welfare of low-skill workers increases, more so when the regulator only puts weight on the welfare of workers.

Table 7: Robustness to Technology Parameters

	Higher ε/σ		Higher ω		Higher ζ	
	utilitarian	only workers	utilitarian	only workers	utilitarian	only workers
Optimal τ_2	0.005	0.015	0.005	0.028	0.008	0.024
Δ Sales share top 1% firms	0.071	0.076	0.070	0.099	0.049	0.059
Δ Sales share corporations	0.115	0.125	0.067	0.096	0.069	0.085
Change in welfare, %	8.09	7.42	5.42	1.52	2.80	2.43
low-skill workers	3.90	6.62	−6.30	1.89	0.20	1.50
high-skill workers	12.0	15.8	−1.84	9.24	8.23	11.8
entrepreneurs	21.8	−6.72	59.2	−8.42	3.74	−12.8

In Table 8 we consider perturbations of the preference and ability parameters. We first increase the relative risk aversion θ to 2, twice as high as in the benchmark calibration. Both the utilitarian regulator, as well as the regulator that seeks to only maximize the welfare of workers, set τ_2 to a value higher than zero, thus increasing product market concentration. As in the benchmark, the utilitarian regulator increases social welfare by mainly benefiting entrepreneurs and high-skill workers. A regulator who only values workers increases the welfare of low-skill workers by 1% and that of high-skill workers by 5.8%, at the expense of entrepreneurs.

While our benchmark calibration matches the Gini coefficients of both wealth and income, as is well-known, absent a fat-tailed distribution of ability shocks, it cannot reproduce the very top wealth and income shares. We show next that our results are robust to matching top wealth and income inequality. We follow Castaneda et al. (2003) and consider an extension with a super-star state. An agent can be in either a normal or a super-star state. In the normal state labor market or entrepreneurial ability follow AR(1) processes with Gaussian shocks. In the super-star state, ability is relatively high. We assume a Markov transition probability between the normal and the super-star state and calibrate these parameters to match the top 1% wealth and income shares.

Table 8 shows that our main findings are robust to this extension: optimal regulation increases product market concentration, more so when the regulator only values the welfare of workers. As in the benchmark, high-skill workers and entrepreneurs are the ones who benefit from optimal utilitarian regulation, whereas low-skill workers benefit as well when the regulator only seeks to maximize worker welfare.

Table 8: Robustness to Preference and Ability Parameters

	Higher θ		Super-Star State	
	utilitarian	only workers	utilitarian	only workers
Optimal τ_2	0.005	0.015	0.010	0.022
Δ Sales share top 1% firms	0.066	0.078	0.074	0.087
Δ Sales share corporations	0.068	0.081	0.071	0.086
Change in welfare, %	3.31	1.49	3.87	2.84
low-skill workers	−1.09	1.00	−0.31	2.79
high-skill workers	2.68	5.83	5.31	9.88
entrepreneurs	56.9	−5.08	24.2	−12.0

4 Conclusions

We study optimal product market interventions in an economy that matches the degree of inequality in the United States and in which firm ownership is highly concentrated and markups increase with firm market share. We proceed in two steps. First, we use a mechanism design approach to characterize optimal regulation in a static setting. Second, we extend the analysis to a richer dynamic setting with capital and wealth accumulation that is more amenable to a quantitative evaluation. Throughout the analysis, we take the general equilibrium and distributional effects of interventions into account.

A robust result that emerges is that optimal regulation nearly restores allocative efficiency and leads to more product market concentration than under the status quo. In addition to increasing aggregate productivity, optimal regulation increases the demand for labor and bids up the equilibrium wages.

We conclude that product market concentration is not costly in and of itself, even in an environment in which firms are owned by a small fraction of households. What is costly is dispersion in the marginal product and wedges that depress the equilibrium wage and the return on capital. Optimal regulation reduces these wedges and in doing so actually increases product market concentration. Our results therefore caution against the widely-held view that reducing concentration and the market power of large firms would necessarily improve the welfare of the poor. Though policies that reduce concentration indeed reduce market power and markups, they have the unintended consequence of also reducing the labor share, aggregate productivity, and wages.

Throughout the paper, we have abstracted from other potentially relevant factors. These include political economy considerations, occupational choice, and externalities through which stifling the growth and entry of smaller, productive firms depresses macroeconomic outcomes. The relevance of these factors for our conclusions depends on how quantitatively important these margins are, for example, on the elasticity of occupational choice to changes in the relative profitability of business versus labor market activities, the extent to which stifling the entry of high-potential firms affects other firms, etc. Measuring the importance of these factors is an important avenue for empirical and theoretical research.

Data Availability Statement The data and code underlying this article are publicly available on Zenodo, at http://doi.org/10.5281/zenodo.11281462.

References

Atkeson, Andrew and Ariel Burstein, "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 2008, *98* (5), 1998–2031.

Atkinson, Antony B., *Inequality: What Can Be Done?*, Harvard University Press, 2015.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen, "Concentrating on the Fall of the Labor Share," *American Economic Review: Papers & Proceedings*, 2017, *107* (5), 180–185.

_ , _ , _ , _ , and _ , "The Fall of the Labor Share and the Rise of Superstar Firms," *The Quarterly Journal of Economics*, 02 2020, *135* (2), 645–709.

Baker, Jonathan and Steven Salop, "Antitrust, Competition Policy, and Inequality," *The Georgetown Law Journal Online*, 2015, *104* (1), 1–28.

Baqaee, David Rezza and Emmanuel Farhi, "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 09 2019, *135* (1), 105–163.

Baron, David P. and Roger B. Myerson, "Regulating a Monopolist with Unknown Costs," *Econometrica*, 1982, *50* (4), 911–930.

Benabou, Roland, "Search, Price Setting and Inflation," *The Review of Economic Studies*, 1988, *55* (3), 353–376.

_ , "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review*, March 2000, *90* (1), 96–129.

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

Benhabib, Jess, Alberto Bisin, and Mi Luo, "Earnings Inequality and Other Determinants of Wealth Inequality," *American Economic Review*, May 2017, *107* (5), 593–97.

Bhandari, Anmol and Ellen McGrattan, "Sweat Equity in U.S. Private Business," *The Quarterly Journal of Economics*, 12 2020, *136* (2), 727–781.

Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz, "Endogenous Entry, Product Variety, and Business Cycles," *Journal of Political Economy*, 2012, *120* (2), 304–345.

_ , _ , and _ , "Monopoly power and endogenous product variety: Distortions and remedies," *American Economic Journal: Macroeconomics*, October 2019, *11* (4), 140–74.

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 25952, National Bureau of Economic Research June 2019.

— **and** — , "Efficient redistribution," *Journal of Monetary Economics*, 2022, *131*, 78–91.

— **and** — , "Markups and Inequality," Working Paper 2022.

Brun, Lidia and Ignacio Gonzalez, "Tobin's Q and Inequality," 2017. working paper.

Burstein, Ariel and Jonathan Vogel, "International Trade, Technology, and the Skill Premium," *Journal of Political Economy*, 2017, *125* (5), 1356–1412.

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.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan, "Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?," *Econometrica*, 2000, *68* (5), 1151–1179.

Colciago, Andrea and Rajssa Mechelli, "Competition and Inequality: Aiyagari meets Bertrand and Cournot," 2019. Working Paper.

da Costa, Carlos E. and Lucas J. Maestri, "Optimal Mirrleesian Taxation in Non-Competitive Labor Markets," *Economic Theory*, November 2019, *68* (4), 845–886.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *The Quarterly Journal of Economics*, 01 2020, *135* (2), 561–644.

DeBacker, J., Vasia Panousi, and Shanthi P. Ramnath, "A Risky Venture: Income Dynamics within the Non-Corporate Private Business Sector," *American Economic Journal: Macroeconomics*, January 2023, *15* (1), 444–74.

Diamond, Peter A, "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, March 1998, *88* (1), 83–95.

Dworczak, Piotr, Scott Duke Kominers, and Mohammad Akbarpour, "Redistribution through Markets," *Econometrica*, 2021, *89* (4), 1665–1698.

Dyrda, Sebastian and Benjamin Pugsley, "Taxes, Regulations of Businesses and Evolution of Income Inequality in the US," 2018. Working Paper.

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

Eeckhout, Jan, Chunyang Fu, Wenjian Li, and Xi Weng, "Optimal Taxation and Market Power," Technical Report, UPF mimeo 2021.

Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski, "Redistribution and Social Insurance," *American Economic Review*, February 2016, *106* (2), 359–86.

Gopinath, Gita and Oleg Itskhoki, "Frequency of Price Adjustment and Pass-Through," *Quarterly Journal of Economics*, 2010, *125* (2), 675–727.

Guner, Nezih, Gustavo Ventura, and Yi Xu, "Macroeconomic Implications of Size-Dependent Policies," *Review of Economic Dynamics*, 2008, *11* (4), 721–744.

Gutiérrez, Germán, Callum Jones, and Thomas Philippon, "Entry costs and aggregate dynamics," *Journal of Monetary Economics*, 2021, *124*, S77–S91.

Hall, Robert E, "New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy," Working Paper 24574, National Bureau of Economic Research May 2018.

_ , "Using Empirical Marginal Cost to Measure Market Power in the US Economy," Working Paper 25251, National Bureau of Economic Research November 2018.

Hartman-Glaser, Barney, Hanno Lustig, and Mindy Xiaolan, "Capital Share Dynamics When Firms Insure Workers," *The Journal of Finance*, 2019, *74* (4), 1707–1751.

Heathcote, Jonathan and Hitoshi Tsujiyama, "Optimal Income Taxation: Mirrlees Meets Ramsey," *Journal of Political Economy*, 2021, *129* (11), 3141–3184.

_ , Fabrizio Perri, Giovanni L. Violante, and Lichen Zhang, "More unequal we stand? Inequality dynamics in the United States, 1967–2021," *Review of Economic Dynamics*, 2023, *50*, 235–266.

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

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.

Hubmer, Joachim, Per Krusell, and Anthony A. Smith Jr, "A Comprehensive Quantitative Theory of the U.S. Wealth Distribution," *NBER Macroeconomics Annual*, 2021, *35*, 391–455.

Jaravel, Xavier and Alan Olivi, "Prices, Non-homotheticities, and Optimal Taxation: The Amplification Channel of Redistribution," Technical Report 2021.

Jones, Charles I., "Intermediate Goods and Weak Links in the Theory of Economic Development," *American Economic Journal: Macroeconomics*, April 2011, *3* (2), 1–28.

Kaplow, Louis, "Market Power and Income Taxation," *American Economic Journal: Economic Policy*, November 2021, *13* (4), 329–54.

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

Kehrig, Matthias and Nicolas Vincent, "Growing Productivity without Growing Wages: The Micro-Level Anatomy of the Aggregate Labor Share Decline," *Quarterly Journal of Economics*, 2021.

Khan, Lina and Sandeep Vaheesan, "Market Power and Inequality: The Antitrust Counterrevolution and its Discontents," *Comparative Political Economy: Regulation eJournal*, 2016.

Kimball, Miles S., "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit, and Banking*, 1995, *27* (4, Part 2), 1241–1277.

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

Kuhn, Moritz and Jose-Victor Rios-Rull, "2013 Update on the U.S. Earnings, Income, and Wealth Distributional Facts: A View from Macroeconomics," *Quarterly Review*, 2016, (April), 1–75.

_ , Moritz Schularick, and Ulrike I. Steins, "Income and Wealth Inequality in America, 1949–2016," *Journal of Political Economy*, 2020, *128* (9), 3469–3519.

Kushnir, Alexey and Robertas Zubrickas, "Optimal Income Taxation with Endogenous Prices," Technical Report 2019.

Pellegrino, Bruno, "Product Differentiation and Oligopoly: a Network Approach," Working Paper 2024.

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 2008, *11* (4), 707–720.

Sachs, Dominik, Aleh Tsyvinski, and Nicolas Werquin, "Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium," *Econometrica*, 2020, *88* (2), 469–493.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *The Review of Economic Studies*, 01 2001, *68* (1), 205–229.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick, "Capitalists in the Twenty-First Century," *The Quarterly Journal of Economics*, 07 2019, *134* (4), 1675–1745.

Stiglitz, Joseph, *The Price of Inequality: How Today's Divided Society Endangers Our Future*, W.W. Norton & Company, 2012.

Waugh, Michael E., *Firm Dynamics and Immigration: The Case of High- Skilled Immigration*, Chicago: University of Chicago Press,