# The Role of Market Power in Taxing Wealth versus Capital Income

Juan Segura
CEMFI

**PRELIMINARY AND INCOMPLETE**

**Abstract:** I study the implications of capital income and wealth taxation in an economy with variable markups and financial frictions. Collateral constraints generate heterogeneous rates of return across entrepreneurs, which lead to significant distinctions between taxing capital income and wealth. Under such frictions, wealth taxation improves allocative efficiency and rises inequality, and the opposite holds for capital income taxation. Motivated by recent empirical evidence on the rise of market power, I revisit these implications in an economy that also includes variable markups and reproduces their dispersion in the US economy. Replacing the capital income tax with a wealth tax while keeping government revenue neutral is not preferable in the presence of variable markups, as aggregate productivity gains are smaller than absent this source of inefficiencies. Productive entrepreneurs that accumulate more resources under wealth taxation and manage to overcome financial frictions also react to the increase in their market power by setting higher prices, obtaining larger profits and not reaching their socially optimal size. Optimal linear taxation asks for a combination of positive tax rates on both wealth and capital income that balances efficiency gains, inequality rises and savings incentives from households.

Keywords: Wealth tax, capital income tax, optimal taxation, redistribution, entrepreneurs, return heterogeneity, misallocation, markups, inequality.

JEL Codes: D43, D60, E62, H21, H24, L13.

# 1 Introduction

The potential benefits of a wealth tax are once again moving up the political agenda. Large economies like the US debate on the proposal of this fiscal instrument, while its presence is unusual in other developed economies — the only OECD countries to use wealth taxation by 2017 were France, the Netherlands, Norway, Spain and Switzerland (OECD, 2018).

Recent evidence on returns heterogeneity (Bach, Calvet and Sodini, 2018; Fagereng et al. 2020) has also reactivated this topic in academia. Taxing capital income is equivalent to taxing wealth in a variety of benchmark economic models,[1] but this equivalence result breaks up under heterogeneous returns to investment, creating significant differences between both fiscal instruments. A key aspect is the source of this heterogeneity. As Guvenen et al. (2023) show in a recent influential paper, in economies with financial frictions, taxing capital income weakens the self-financing channel of productive entrepreneurs by reducing after-tax profit gains, limiting their financing capacity and increasing aggregate misallocation.[2] On the contrary, wealth taxation expands the tax base and shifts part of the tax burden towards less productive entrepreneurs, favouring wealth accumulation of efficient producers and helping them overcome credit limitations. But while the efficiency implications for taxing capital income or wealth are clear in economies with financial constraints, heterogeneity in returns may also be driven by other reasons.

I argue an important source of returns heterogeneity is the presence of product market power and the existence of heterogeneous markups across firms. The work of De Loecker, Eeckhout and Unger (2020) document an increase of 40 percentage points in aggregate markups in the last 50 years in the US. More importantly, this aggregate increase has been mostly driven by a fatter upper tail of the markup distribution: there has been a three-fold increase in the difference between the 90-th percentile and the median of the markup distribution during the same period of time — while the median has not changed. The large market shares owned by some businesses allow them to charge higher markups, produce at suboptimal levels (from a social planner perspective) and extract oligopoly rents, becoming an important source of heterogeneity in returns. This has implications for resource allocation and welfare, which has obvious ramifications for redistribution policies.

This project attempts to answer the question: what is the optimal combination of wealth and capital income taxation in an economy combining financial frictions and heterogeneity in markups? For this purpose, I build a model of firm dynamics with both sources of heterogeneous returns, analyze the inefficiencies created by each source and their interaction, and numerically solve the problem of a social planner that aims to maximize welfare given the fiscal instruments at hand.

The taxation exercise shows the welfare-maximizing policy involves positive tax rates on both

[1] Equivalence in this context means it is possible to find a tax rate on wealth such that tax government revenue and after-tax financial income for all households are the same as under a tax schedule taxing capital income instead of wealth.

[2] In an economy like the one in their paper, heterogeneity in returns is driven by differences in productivities (as more efficient entrepreneurs obtain larger profits every thing else equal) and wealth (as low wealth entrepreneurs' firms might not be able to reach the sizes that maximize their profits).

wealth and capital income. Low rates of capital income taxation (together with high rates of wealth taxation to preserve government revenue levels) partially undo the misallocation generated by financial frictions. However, fiscal reforms of this nature cannot correct the inefficiencies coming from market power, which, if anything, worsen. The productive entrepreneurs that accumulate more wealth and escape financial limitations under a tax on wealth exploit the acquired market power by charging higher markups and not reaching their optimal size, thus reducing aggregate efficiency gains. These same entrepreneurs also generate larger profits, leading to a more unequal economy. Thus, the social planner must balance efficiency, equity and capital accumulation responses. Having to choose between taxing only capital income or wealth, a social planner would prefer the former (the status quo), as otherwise redistribution losses would swamp the drop in misallocation. Indeed, the optimal policy is a small deviation from the benchmark economy where only capital income is taxed, although it asks for some degree of wealth taxation. This differs from previous research where financial frictions is the only source of misallocation.

The economic environment consists of households that derive utility from consumption and supply labor to a final good producer. Households differ in their labor productivity, which follows a persistent stochastic process. Moreover, a fraction of them also posses an entrepreneurial productivity following a persistent stochastic process that allows them to run their own businesses. Households can partially insure against risk through saving in a bond market. Each entrepreneur can produce a differentiated intermediate good through a private technology, and the same bond market used for savings allows producers borrowing to invest in their own firm above their own assets but up to a collateral constraint. The different intermediate goods are combined in a Kimball (1995) aggregator by the final good producer, which defines the demand for these varieties. Under this demand, entrepreneurs find optimal to charge a price equal to a markup over their marginal cost, where the markup increases with their quantity market share. The government finances its expenditure through tax revenues.

Given the focus on efficiency and redistribution, the model is calibrated to match some salient features of the US wealth and income distributions, the dispersion in the distribution of markups, and some key aggregate ratios.

The reminder of the paper is the following. Section 2 lays out the model economy. Section 3 elaborates on the inefficiencies caused by financial constraints and markups. Section 4 describes the calibration procedure. Section 5 presents the optimal policy results. Finally, section 6 concludes.

## Related Work

This project relates to the literature on optimal taxation under heterogeneous returns, and more broadly to the work on optimal capital income taxation (Conesa, Kitao and Krueger, 2009). The point of departure is Guvenen et al. (2023), who study optimal capital income and wealth taxation in a life-cycle economy with financial frictions. In comparison, my model includes market power and heterogeneous markups on top of financial frictions, although the rest of the production side is kept purposefully similar to theirs in order to better identify the effect of variable markups on the efficiency implications for each tax. The presence of variable markups reduce aggregate efficiency gains from taxing wealth as opposed to taxing capital income, and when calibrating the model to match key inequality and market power statistics, results differ from those of Guvenen et al. (2023): moving from taxing capital income to taxing wealth is welfare decreasing, and optimal taxation rates are a modest deviation from the benchmark economy. Conversely, this is more aligned with the findings of Boar and Midrigan (2023). They show that in an economy where private firms coexist with unconstrained corporate firms, the motive for redistribution is larger than the losses from misallocation, making optimal capital income taxation preferred to optimal wealth taxation. This is because public firms do not operate under rigid financial limitations, implying a lower level of aggregate inefficiencies and smaller need for reallocation. Boar and Knowles (2024) study the importance of endogenizing financial frictions in this debate. Their results align with mine in that optimal taxation sets positive tax rates on both capital income and wealth. Also related is the work of Macnamara, Pidkuyko and Rossi (2022), who find that optimal policy calls for not taxing capital income nor wealth as long as both consumption taxes and the progressivity of the labor income tax schedule can be increased. This work attempts to expand this body of literature by revisiting the debate between capital income and wealth taxation while accounting for market power inefficiencies.

Finally, this study shares points of contact with other strands of the literature. It has obvious connections with papers studying entrepreneurial firms with differences in productivity facing credit limitations, mostly pioneered by Quadranini (2000) and later followed by Cagetti and De Nardi (2009), Buera, Kaboski and Shin (2011), Midrigan and Xu (2014) and Moll (2014), among others. It also relates to recent papers looking at optimal taxation in environments with entrepreneurs (Imrohoroglu, Kumru and Nakornthab, 2018; Brüggemann, 2021). In parallel, it also touches upon the literature looking at the aggregate and distributional costs of market power and variable markups. Some examples go from the seminal studies of Dixit and Stiglitz (1977) and Atkeson and Burstein (2008), to the more recent work of Boar and Midrigan (2019), Baqaee and Farhi (2020) and Edmond, Midrigan and Xu (2023), among others.

# 2 Model

I present an incomplete-markets model of monopolistic competition with variable markups and financial frictions. All households are endowed with some labor productivity and inelastically supply it to the labor market.[3] Moreover, a fraction of households also have an entrepreneurial productivity and can run firms that produce varieties of an intermediate good. There is a perfectly competitive final good producer that uses labor and efficiency-adjusted capital as inputs, where the latter is an aggregate of the varieties generated by entrepreneurs. Last, there is a government that finances its spending through linear taxes.

[3]It can instead be interpreted that households are given labor endowments that they supply inelastically.

## 2.1 Final Good Producer

The final good is produced according to a Cobb-Douglas technology:

$$Y = Q^{\alpha} L^{1-\alpha},$$

where $L$ is aggregate labor and $Q$ is an efficiency-adjusted capital stock combining firm-specific inputs $q_i$ through a Kimball (1995) aggregator implicitly defined by

$$\int_0^{\psi} \Upsilon\Big(\frac{q_i}{Q}\Big) di = 1,$$

where $\Upsilon(1) = 1, \Upsilon' > 0$, $\Upsilon'' < 0$, and $\psi \in (0, 1]$ is the fraction of households that can engage in production.[4] Anticipating the calibration of the model, I follow Klenow and Willis (2016) and assume an aggregator of the form

$$\Upsilon(x) = 1 + (\sigma - 1) \exp\Big(\frac{1}{\varepsilon}\Big) \varepsilon^{\sigma/\varepsilon - 1} \Bigg(\Gamma\Big(\frac{\sigma}{\varepsilon}, \frac{1}{\varepsilon}\Big) - \Gamma\Big(\frac{\sigma}{\varepsilon}, \frac{x^{\varepsilon/\sigma}}{\varepsilon}\Big)\Bigg),$$

where $\Gamma(u, z)$ is the upper incomplete gamma function, $\Gamma(u, z) = \int_z^{\infty} s^{u-1} e^{-s} ds$. Under this specification the price elasticity of demand is given by

$$\epsilon(x) = -\frac{\Upsilon'(x)}{x\Upsilon''(x)} = \sigma x^{-\frac{\varepsilon}{\sigma}}. \tag{1}$$

Thus, if $\varepsilon > 0$, entrepreneurs with larger quantity market shares $x = q/Q$ face less elastic demand curves, allowing them to charge higher markups.

This aggregator also produces a non-zero superelasticity, the elasticity of the price elasticity with respect to relative size,

$$-\frac{d\epsilon(x)}{dx}\frac{x}{\epsilon(x)} = \frac{\varepsilon}{\sigma}.$$

The Kimball aggregator is often seen as a generalization of the familiar constant elasticity of substitution (CES) aggregator, which is obtained in the limit case where $\varepsilon \to 0$.

The demand structure generated by the Kimball aggregator generates a negative relationship between market shares and demand elasticities and generates variable markups, like in models of oligopolistic competition (Atkeson and Burstein, 2008). [5]

[4]The subscript in the aggregator refers to the firm $i$ in the continuum of firms with measure $\psi$ (it is the same continuum as that of households with entrepreneurial productivity endowments that engage in production). As we will see, households in the model are distributed across the state space on productivities and wealth, so when convenient, I will identify individuals' choices by a triple $(s, z, a)$ rather than the index $i$.

[5]However, with Kimball aggregation firms are atomistic as under monopolistic competition, thus sweeping any strategic interaction between individual entrepreneurs and simplifying the computational solution of the problem.

The final good producer maximizes profits by choosing the optimal combination of intermediate good varieties and labor, subject to the technology that aggregates these varieties,

$$\max_{\{q_i\},L} \; Q^\alpha L^{1-\alpha} - \int_0^\psi p_i q_i di - wL \qquad \text{s.t.} \quad \int_0^\psi \Upsilon\Big(\frac{q_i}{Q}\Big) dj = 1,$$

where $p_i$ is firm i's intermediate good price and $w$ is the equilibrium wage. The solution gives standard demand functions[6]

$$Q = \alpha \frac{Y}{p_i} \Upsilon'\Big(\frac{q_i}{Q}\Big) D, \qquad L = (1-\alpha)\frac{Y}{w},$$

where $D = \Big[ \int_0^\psi \Upsilon'\Big(\frac{q_i}{Q}\Big) \frac{q_i}{Q} di \Big]^{-1}$ is an endogenously determined demand index.[7]

## 2.2 Households

The economy is populated by a continuum of households (with measure normalized to one) that derive utility from consumption $c$ of a final good and inelastically supply one unit of labor to the final good producer. Households share the same preferences and discount rate $\beta$ but differ in their labor productivity, $s \in \boldsymbol{S}$, assumed to follow a stochastic Markovian process $G_s(s'|s)$. In addition, a fraction $\bar{\psi}$ of households are also endowed with an entrepreneurial productivity, $z \in \boldsymbol{Z}$, also Markovian, $G_z(z'|z)$.[8] Their optimization problem can be split into its dynamic and static parts.

### 2.2.1 Dynamic problem

Each household maximizes its discounted lifetime utility with respect to consumption and savings subject to a budget constraint and a no borrowing condition:[9]

$$\begin{aligned} V(s,z,a) &= \max_{c,a'} \; u(c) + \beta\, \mathbb{E}\Big[V(s',z',a')\Big] \\ \text{s.t.} \quad & (1+\tau_c)c + a' = (1-\tau_l)ws + \omega(z,a;\tau_a,\tau_k) \\ & a' \geq 0, \end{aligned}$$

where $r$ is the return on savings, $w$ is the labor wage and $\{\tau_c, \tau_l, \tau_a, \tau_k\}$ is the set of flat-rates for taxes on consumption, labor income, wealth and capital income, respectively. The object $\omega$ is the

Moreover, despite the absence of microfoundations leading to variable markups like in models of oligopolistic competition, both market structures deliver the same qualitative results (see Edmond, Midrigan and Xu, 2023).

[6]The derivation of demands is standard and can be found in section A.1.1 of the appendix.

[7]The index $D$ summarizes all relevant information about competing firms.

[8]For simplicity, it is assumed that productivity processes are independent.

[9]To ease notation, it is safe to assume the share of households with no entrepreneurial productivity endowments have $z = 0$ at all times.

after-production and after-tax non-labor income of the household,

$$\omega(z, a; \tau_a, \tau_k) = (1 - \tau_a)a + (1 - \tau_k)[ra + \pi(z, a)].$$

The intertemporal solution to this problem is characterized by the Euler equation

$$u_c(c) \geq \beta\, \mathbb{E}\left[\omega_a(z', a') \cdot u_c(c')\right],$$

which together with the budget constraint defines the policy functions on savings and consumption $g_a(s, z, a)$ and $g_c(s, z, a)$, respectively.

### 2.2.2 Static problem

Intermediate goods are produced by entrepreneurs with access to a backyard technology linear in capital $q = zk$. They face a collateral constraint that limits their borrowing level to a multiple $\lambda \geq 1$ of their wealth. In theory, they also compete with other firms, but since there is a continuum of them, individual firms are atomistic and any strategic interactions between them are captured by aggregate demand indexes. They take as given beliefs about aggregate quantities and input prices. Thus, the profit maximization problem is

$$\begin{aligned} \pi(z, a) = \max_{p,q,k} \; & \Big\{ p(q; Q, D)q - (r + \delta)k \Big\} \\ \text{s.t.} \quad & q = zk \\ & p = \alpha \frac{Y}{Q} \Upsilon'\Big(\frac{q}{Q}\Big) D = \Upsilon'\Big(\frac{q}{Q}\Big) X \\ & k \leq \lambda a. \end{aligned}$$

The solution to this problem is

$$p = \underbrace{\frac{\sigma}{\sigma - (q/Q)^{\varepsilon/\sigma}}}_{\text{Markup}} \Big( \underbrace{\frac{r + \delta + \Lambda}{z}}_{\text{Marginal cost}} \Big) \tag{2}$$

and $\Lambda(q - z\lambda a) = 0$, where $\delta$ is the depreciation rate and $\Lambda$ is the Lagrange multiplier corresponding to the collateral constraint. Entrepreneurs charge a price equal to a markup, increasing in the relative quantity they produce, over a marginal cost that includes the cost of capital and the cost of satisfying the credit constraint. Quantities can be found by equating firms' FOCs with the final good producer inverse demand function,

$$\underbrace{\Upsilon'\Big(\frac{q^*}{Q}\Big) X}_{\text{Demand curve}} = \underbrace{\frac{\sigma}{\sigma - (q^*/Q)^{\varepsilon/\sigma}} \Big( \frac{r + \delta + \Lambda}{z} \Big)}_{\text{Supply curve}}, \tag{3}$$

yielding the solution

$$q(z,a) = \min\Big\{q^*(z), z\lambda a\Big\} \iff g_k(z,a) = \min\Big\{q^*(z)/z, \lambda a\Big\},$$

where $g_k(z,a)$ is the optimal capital choice function. The associated maximized profit function is

$$\pi(z,a) = \begin{cases} q^*(z)\Big[\Upsilon'\Big(\frac{q^*(z)}{Q}\Big)X - \frac{(r+\delta)}{z}\Big] & \text{if} \quad k(z,a) \neq \lambda a \\ z\lambda a\Big[\Upsilon'\Big(\frac{z\lambda a}{Q}\Big)X - \frac{(r+\delta)}{z}\Big] & \text{if} \quad k(z,a) = \lambda a \end{cases}.$$

Note not all firms decide to engage in production. There is a positive (strictly, if we rule out the case where nobody produces) price above which no intermediate goods are demanded: $p(q=0) = \alpha\frac{Y}{Q}\frac{\sigma-1}{\sigma}e^{1/\varepsilon}D$. It follows only efficient enough households will decide to become entrepreneurs. This cutoff solves (3) for $q=0$ and, provided $a>0$, is given by

$$\bar{z} = (r+\delta)\frac{1}{\alpha}\frac{Q}{Y}\frac{\sigma}{\sigma-1}e^{-1/\varepsilon}\frac{1}{D}.$$

Entrepreneurs with $z \leq \bar{z}$ would find optimal to produce negative quantities and will not engage in production. On the contrary, those with $z > \bar{z}$ will run their firms and, if credit constrained, will produce as much as possible given limitations, which will always be a positive quantity with a price below the choke price defining the cutoff $\bar{z}$.

## 2.3 Government

The government finances some exogenous spending G through taxes on consumption, labor income, wealth and capital income at flat rates $\{\tau_c, \tau_l, \tau_a, \tau_k\}$, respectively. For simplicity, it is assumed government spending provides no utility to households.

## 2.4 Equilibrium

A steady state equilibrium is a set of decision rules $\{g_c(s,z,a), g_a(s,z,a), g_k(s,z,a)\}$, a pair of aggregates $\{Q, L\}$, a set of prices $\{p(x), r, w\}$, a set of taxes $\{\tau_c, \tau_l, \tau_k, \tau_k\}$, and a probability measure of individuals $\Gamma(s,z,a)$ such that

1. The functions $\{g_c(s,z,a), g_a(s,z,a), g_k(s,z,a)\}$ maximize the discounted utility of households given prices and taxes.

2. The solution to the final good producer's problem sets the pricing of varieties function $p(x)$ and wages $w$.

3. Production factors come from aggregation of households and firms decisions:

$$\int_{s,z,a} \Upsilon\Big(\frac{z \cdot g_k(s,z,a)}{Q}\Big) d\Gamma(s,z,a) = 1, \quad L = \int_{s,z,a} s \cdot d\Gamma(s,z,a).$$

4. The capital market clears:

$$\int_{s,z,a} (a - g_k(s,z,a)) d\Gamma(s,z,a) = 0.$$

5. The government budget balances:

$$G = \tau_c \int_{s,z,a} g_c(s,z,a) d\Gamma(s,z,a) + \tau_l w L + \tau_k \int_{s,z,a} (ra + \pi(z,a)) d\Gamma(s,z,a) + \tau_a \int_{s,z,a} a \cdot d\Gamma(s,z,a).$$

# 3 Cost of Financial Constraints and Markups

The baseline economy is distorted due to the presence of financial constraints in the form of collateral and variable markups across different variety producers. This section shows how these two distortions generate misallocation at the individual level, alter aggregate input choices and decrease aggregate productivity.

## 3.1 Aggregates Losses

Recall the equilibrium condition for the market of intermediate good varieties (3) states the price being equal to the product of a markup and a marginal cost including the potential cost of satisfying the collateral. In the spirit of the literature framing the discussion of misallocation as wedges (Hsieh and Klenow, 2009), it is possible to express the multiplier coming from the collateral as a wedge on capital cost,

$$r + \delta + \Lambda = (r + \delta)(1 + \tau^k) \Rightarrow \tau^k = \frac{\Lambda}{r + \delta},$$

so the optimality condition (2) can be rewritten as

$$p(x(z,a)) = \mu(x(z,a)) \frac{(r + \delta)(1 + \tau^k(z,a))}{z}.$$

Absent financial frictions and variable markups, average revenue product of capital (ARPK) is equalized across variety producers. Thus, the presence of these two terms in the expression for ARPK points out at the misallocation they generate

$$\text{ARPK} \equiv \frac{p(x(z,a)) q(z,a)}{k(z,a)} = \mu(x(z,a))(1 + \tau^k(z,a))(r + \delta).$$

Motivated by this expression, define the aggregate demand of capital as $K = \int_{z,a} g_k(z,a) d\Gamma(z,a)$, let $\mathcal{M}$ denote the aggregate markup of the economy and $1+\mathcal{T}^k$ the aggregate capital wedge, implicitly defined by

$$\frac{P_Q Q}{K} = \mathcal{M}(1+\mathcal{T}^k)(r+\delta),$$

where $P_Q$ is the price of efficiency-adjusted capital Q.[10,11] Similarly to the expression for an individual firm, aggregate ARPK shows how the aggregate markup and capital wedge lead to suboptimal capital accumulation.

## 3.2 Dispersion and Productivity

Dispersion in markups and wedges leads to differences in ARPK and reduces aggregate productivity. To see this, consider the problem of a social planner that aims to maximize aggregate output Y subject to a technology defined by the Kimball aggregator and a resource constraint on total available capital:

$$\begin{aligned}
&\max_{q(z,a),Q} Y \\
\text{s.t.} \quad & Y = Q^{\alpha} L^{1-\alpha} \\
& \int_{z,a} \Upsilon\Big(\frac{q(z,a)}{Q}\Big) d\Gamma(z,a) = 1 \\
& \int_{z,a} \underbrace{\frac{q(z,a)}{z}}_{k(z,a)} d\Gamma(z,a) = K.
\end{aligned}$$

The first order conditions are

$$\begin{aligned}
&\alpha \frac{Y}{Q} = \upsilon \int_{z,a} \Upsilon'\Big(\frac{q(z,a)}{Q}\Big) \frac{q(z,a)}{Q^2} d\Gamma(z,a) \Rightarrow \upsilon = \alpha \frac{Y}{D}, \\
&\upsilon \Upsilon'\Big(\frac{q(z,a)}{Q}\Big) \frac{1}{Q} - \frac{\nu}{z} = 0,
\end{aligned}$$

where $\upsilon$ and $\nu$ are the multipliers coming from the technology and resource constraints, respectively. Combining the two conditions, we obtain the relationship defining the efficient allocation of resources across entrepreneurs,

$$\Upsilon'\Big(\underbrace{\frac{q(z)}{Q}}_{x(z)}\Big) X = \frac{\nu}{z},$$

[10] It follows $P_Q Q = \int_{s,z,a} p(z,a) q(z,a) d\Gamma(s,z,a)$.

[11] See section A.1.3 in the appendix for a derivation of the aggregate markup and the aggregate capital wedge.

which only depends on the productivity level of workers and not their wealth level.[12] The efficient allocation states the familiar optimality condition that prices must equal marginal costs (up to a multiplicative constant). On the contrary, entrepreneurs seek profit maximization and are subject to credit constraints, so they equate prices to a markup over a marginal cost potentially inflated by credit limitations:

$$\Upsilon'(x(z,a))X = \mu(x(z,a))\frac{(1+\tau^k(z,a))(r+\delta)}{z}.$$

Thus, variable markups and capital wedges distort optimal production quantities and decrease aggregate productivity.

# 4 Quantification

The parameterization strategy sets a subset of parameters to assigned conventional values, while the remaining ones are calibrated to match salient features of the wealth and income US distributions, as well as the dispersion in markups. All parameters except the tax rates are kept constant through the quantitative exercises.

## 4.1 Assigned Parameters

One time period represents one year. The model assumes an isoelastic utility function, $u(c) = (c^{1-\gamma}-1)/(1-\gamma)$ with $\gamma = 1.5$. The depreciation rate of capital is set to $\delta = 0.06$. The elasticity of efficiency-adjusted capital is set to $\alpha = 0.36$ to match an aggregate labor share of 0.64. The US tax schedule is represented by a triplet of flat-tax rates $(\tau_c, \tau_l, \tau_k)$. Following Guvenen et al. (2023), I set $\tau_c = 7.5\%$, $\tau_l = 22.4\%$, $\tau_k = 25.0\%$. The tax rate on wealth is zero. All these choices are summarized in Panel A from Table 1.

## 4.2 Calibrated Parameters

Individual labor endowments are assumed to follow an AR(1) process in logs with persistence $\rho_s$ and innovations with standard deviation $\theta_s$. Similarly, individual entrepreneurial productivity is assumed to follow an AR(1) process in logs with persistence $\rho_z$ and standard deviation $\theta_z$. Both processes are assumed to be independent of each other.

The fraction of households with a strictly positive entrepreneurial productivity draw, the discount factor, the strength of the collateral constraint, the parameters governing labor endowment and entrepreneurial productivity, and the parameters shaping the Kimball aggregator are selected to

[12] Trivially, this follows from the lack of collateral. Solving the planner's problem with a credit constraint gives $\Upsilon'\left(\frac{q}{Q}\right)X = \frac{\nu+\Lambda}{z}$, which is closer to the individual entrepreneur's optimality condition. Although not including financial constraints may seem an unfair comparison to the laissez faire solution, the goal of this section is to exhibit the inefficiencies caused by both markups and financial constraints.

Table 1: Parameterization

| | Panel A. Assigned parameters | |
|---|---|---|
| $\gamma$ | Relative risk aversion | 1.5 |
| $\alpha$ | Capital elasticity | 0.36 |
| $\delta$ | Capital depreciation rate | 0.06 |
| $\tau_c$ | Consumption tax rate | 7.5% |
| $\tau_l$ | Labor income tax rate | 22.4% |
| $\tau_k$ | Capital income tax rate | 25.0% |
| | *Panel B. Calibrated parameters* | |
| $\bar{\psi}$ | Fraction of households with entrepreneurial productivity | 0.356 |
| $\beta$ | Discount factor | 0.952 |
| $\lambda$ | Collateral strength | 1.79 |
| $\rho_s$ | Persistence for labor endowment | 0.765 |
| $\theta_s$ | Volatility of labor endowment innovations | 0.575 |
| $\rho_z$ | Persistence for entrepreneurial productivity | 0.968 |
| $\theta_z$ | Volatility of entrepreneurial productivity innovations | 0.369 |
| $\sigma$ | Average demand elasticity | 9.89 |
| $\varepsilon/\sigma$ | Super-elasticity of demand | 0.30 |

Notes: flat-tax rates are obtained from Guvenen et al. (2023). The rest of parameters in Panel A take standard values. Parameters in Panel B are jointly internally calibrated.

target the moments reported in Table 2. The selected statistics are the fraction of entrepreneurs reported by Boar and Midrigan (2023), a capital-to-output ratio of 3.00, the business-debt-to-GDP ratio reported by Guvenen et al. (2023), gini indices and various percentiles from the wealth and income distributions computed by Kuhn and Ríos-Rull (2020), and interpercentile differences at the top of the (sales-weighted) markup distribution from De Loecker, Eeckhout and Unger (2020).

As reported in Table 2, the model does a good job reproducing the targeted moments. The fraction of entrepreneurs in the economy and the two selected aggregate ratios perfectly match the values reported in the literature. The wealth and income distributions are successfully matched, specially at the 50%, 10% and 1% percentile levels. High levels of wealth concentration at the top are driven by the profits of entrepreneurs and their increased incentives to save to escape financial frictions. Last, the gaps between the 90th and the 50th and 75th percentiles of the markup distribution show considerable dispersion in product market power between firms at the high-end and the median. The average percentage deviation between targets and the benchmark calibration is 3.4%.

Table 2: Targeted moments

| Target | US Data | Model |
|---|---|---|
| | *Aggregates* | |
| Fraction entrepreneurs | 0.117 | 0.117 |
| K/Y | 3.00 | 3.00 |
| Business Debt/GDP | 1.29 | 1.29 |
| | *Inequality* | |
| Gini wealth | 0.85 | 0.79 |
| Top 50% wealth share | 0.985 | 0.985 |
| Top 10% wealth share | 0.764 | 0.749 |
| Top 1% wealth share | 0.372 | 0.358 |
| Gini income | 0.57 | 0.55 |
| Top 50% income share | 0.855 | 0.871 |
| Top 10% income share | 0.467 | 0.459 |
| Top 1% income share | 0.192 | 0.188 |
| | *Markups dispersion* | |
| p90-p50 markup | 1.30 | 1.37 |
| p90-p75 markup | 0.90 | 0.86 |

Notes: the fraction of entrepreneurs is taken from Boar and Midrigan (2023). Ratio between business debt and GDP follows from Guvenen et al. (2023). Wealth and income inequality measures are obtained from Kuhn and Ríos-Rull (2020). Dispersion in the sales-weighted markup distribution follows from De Loecker, Eeckhout and Unger (2020).

Panel B of Table 1 report the calibrated parameters. The fraction of household with strictly positive entrepreneurial productivity is 0.356. This number does not necessarily have to match the fraction of entrepreneurs in the economy due to the presence of a choke price described above. The discount factor is $\beta = 0.952$ and the maximum leverage ratio is $\lambda = 1.79$. The persistence of the labor endowment is $\rho_s = 0.765$ and the standard deviation of its innovations is $\theta_s = 0.575$. The entrepreneurial productivity process is more persistent and less volatile, with persistence and standard deviation of innovations values of $\rho_z = 0.968$ and $\theta_z = 0.369$, respectively. The demand elasticity of a firm with relative size $x = 1$ is $\sigma = 9.89$, and the super-elasticity of demand is $\varepsilon/\sigma = 0.30$.

Despite all parameters being jointly calibrated, some are more tightly linked to certain targets. Intuitively, the discount factor is pinned down by the capital-to-output ratio, and the maximum leverage ratio by the aggregate level of business debt. The moments from the wealth and income distributions help to pin down the parameters governing the labor and entrepreneurial ability processes. Last, the parameters from the Kimball aggregator rule the responsiveness of market power to firm size and are pinned-down by the dispersion in markups. However, productivity parameters significantly affect targets originally linked to market power parameters, and vice versa.

It is worth discussing the implications of the benchmark calibration for the econmy distortions. The model predicts large levels of misallocation, with the aggregate TFP being 73% lower than absent variable markups and financial constraints. Moreover, the aggregate markup equals 1.72, larger than common values in the literature. This follows from the calibration targeting the dispersion in markups, which requires high sensitivity of markups to relative sizes. The choice of targeting markups dispersion instead of aggregate levels is motivated by the former being more tightly linked to misallocation in an economy (see Hsieh and Klenow, 2009). Still, the low maximum leverage level required to match aggregate relative debt levels imply higher aggregate distortions and misallocation coming from financial constraints. In particular, the implied aggregate capital wedge is $1 + \mathcal{T}^k = 2.06$ and the standard deviation of individual capital wedges is $sd(1 + \tau^k) = 1.67$, much larger than the standard deviation of markups $sd(\mu) = 0.33$.

# 5 Optimal Taxation

This section studies the combination of wealth and capital income taxes that maximizes welfare. First, I discuss the effects of assets accumulation by entrepreneurs on the different sources of inefficiencies. Then, I define the concept of neutrality that ensures comparability across tax systems and define a measure of welfare. Last, I discuss the results. All comparisons are between steady-states and the planner is limited to use flat-rate taxes.

## 5.1 Wealth accumulation and Wedges

Increasing the wealth stock level of entrepreneurs can have important aggregate efficiency consequences. Imagine we could transfer resources to an entrepreneur. If the entrepreneur has enough resources before the transfer so she is producing at her desired level (i.e. the financial constraint is not binding), her production will not change with the transfer. On the contrary, transferring resources to an entrepreneur for whom the financial constraint binds will reduce distortions coming from collateral and increase her firm size. However, in an economy with both financial constraints and markups sensitive to size, increasing the wealth level of an entrepreneur will have side-effects on her production decisions, as summarized in the following proposition.

**Proposition 1.** If $\varepsilon > 0$, an increase in the wealth stock of an entrepreneur leads to a weakly positive change in the markup it charges over its marginal cost,

$$\frac{\mathrm{d}\mu}{\mathrm{d}a} \geq 0.$$

*Proof:* see Appendix A.1.5.

This result states that, in markets where markups increase with relative firm-size[13] and under the presence of collateral, increasing the wealth level of entrepreneurs will weakly increase their markups (strongly increase if the financial constraint binds). Intuitively, if a constrained entrepreneur increases her wealth level, she will be able to increase the level of capital that goes into production and scale up her business. This growth implies her firm is now relatively larger than before, providing her with a more dominant position in the market that she can exploit by charging larger markups.

## 5.2 Neutrality

To compute the optimal policy, it is necessary to take a stand on the concept of neutrality. This paper imposes revenue neutrality to guarantee comparability across steady states: for a given capital income tax, the tax rate on wealth is pinned down by making the government revenue the same as in the benchmark economy.

## 5.3 Welfare

This paper takes a Benthamite stand to quantify the welfare consequences of different tax rates combinations. Formally, let $v_b(s, z, a, \Delta)$ be the equilibrium value function before a tax change of a household in state $(s, z, a)$, whose consumption allocation is changed by a percentage $\Delta$ in present and future time points,

$$v_b(s, z, a, \Delta) = u(c_b(s, z, a)(1 + \Delta)) + \beta\, \mathbb{E}\left[v_b(s', z', a', \Delta)\right],$$

where $c_b(s, z, a)$ is the consumption policy function before the tax change evaluated at $(s, z, a)$. The welfare gain or loss from the tax change is defined as the percentage of additional consumption $\Delta_{\text{tr}}$ that should be given to all households in the benchmark economy so the aggregate welfare is the same as in the new economy. Formally, $\Delta_{\text{tr}}$ solves

$$\int_{s,z,a} v_b(s, z, a, \Delta_{\text{tr}}) d\Gamma_b(s, z, a) = \int_{s,z,a} v_a(s, z, a, 0) d\Gamma_a(s, z, a),$$

where $v_a(s, z, a, \Delta)$ is the value function in the new economy, and $\Gamma_b(s, z, a)$ and $\Gamma_a(s, z, a)$ are the before and after tax equilibrium distributions of agents, respectively.

## 5.4 Results

Figure 1 plots how different economic measures change as the economy moves from the benchmark capital income tax rate of 25.0% (and the wealth tax rates adjusts accordingly). A social planner that aims to maximize welfare in this economy through a combination of wealth and capital income

[13] This is an extended assumption in models with heterogeneous markups, both under monopolistic and oligopolistic competition.

taxes while keeping government revenue neutral would tax wealth at 0.56% and capital income at 18% (this corresponds to the left-most vertical dashed line in the panels from Figure 1).[14] That is, optimal taxation asks for a combination of positive tax rates on both wealth and capital income. The steady-state under the new tax rates implies a welfare increase equivalent to a 0.2% increase in consumption for all agents. The welfare increase between the optimum and benchmark economies may seem small because both economies are close. Next, I discuss how the optimal steady-state balances efficiency gains, inequality and savings incentives from households.

The top panels of the figure show how the wealth tax rate adjusts to achieve revenue neutrality as we move from the benchmark economy and the associated aggregate distortions. As expected, there is a negative relationship between both tax rates since one is the margin of adjustment for the other. Taxing wealth rather than capital income strengthens the self-financing channel of wealth-poor efficient entrepreneurs, allowing them to accumulate wealth more rapidly and partially or fully escape from financial limitations. On the contrary, taxing capital income weakens their primary channel to accumulate wealth. This is the mechanism discussed in Guvenen et al. (2023) and it explains the increasing profile in the aggregate capital wedge panel. However, as described in Proposition 1, in an economy with collateral and variable markups that increase with the relative size of the firm, a drop in the individual capital wedge coming from the financial frictions leads to an increase in the individual markup. Thus, the social planner has to balance both sources of inefficiencies.

The middle row panels illustrate the degree of misallocation in the economy, the (asset-weighted) marginal returns to savings, and the associated response of output. Under the present calibration, capital wedges are much more responsive than markups to the different tax combinations. This explains why misallocation is also monotonic increasing in $\tau_k$. Even though markups are much less sensitive to tax changes than capital wedges, they still play an important role in this problem. In contrast to an scenario with only financial frictions, having another source of inefficiencies reduces the effectiveness of wealth taxation (as opposed to capital income taxation) in reducing aggregate wedges and total misallocation. Marginal returns to savings, here defined as the derivative of non-labor income with respect to assets $\omega_a(z,a) = (1-\tau_a) + (r + \pi_a(z,a))(1-\tau_k)$, are an increasing function of $\tau_k$ (this holds for both workers and entrepreneurs). For workers, this is entirely driven by general equilibrium effects defining how interest rates adjust for tax changes.[15] On top of that, entrepreneurs have larger incentives to save in order to escape financial constraints that, as we will see, become more binding as capital income tax rates increase.

The combination of misallocation and marginal returns to savings define the efficiency of the economy and its degree of capital accumulation, thus explaining the behaviour of aggregate output. The optimal economy is 1% larger than the benchmark, which is only half from the 2% increase achieved when capital income is taxed at 0.5%. Despite the economy not being at its largest capacity,

[14] Without taking into account transitions between steady states, which are expected to be part of the analyses in future versions of the paper.

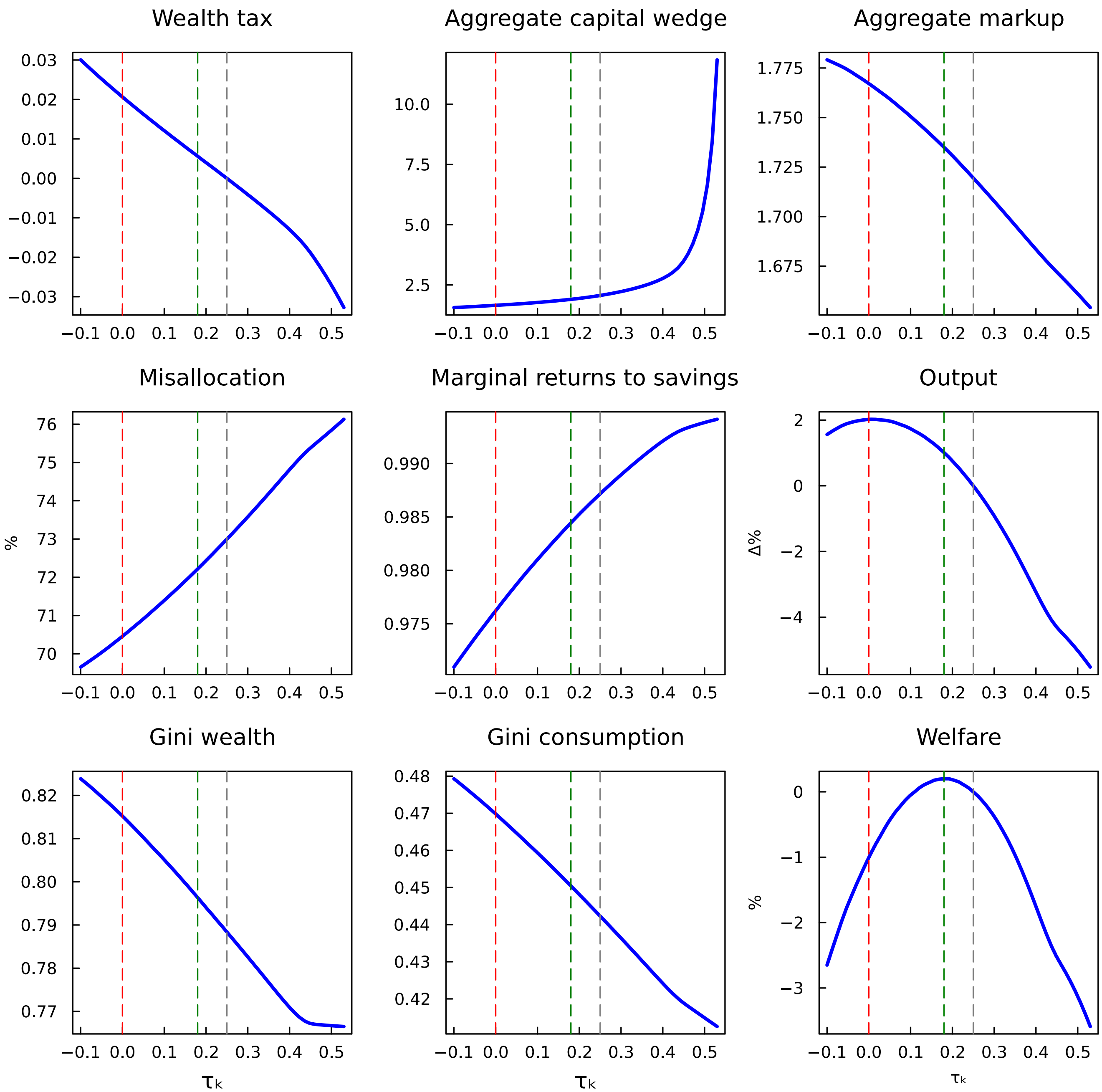
[15] The marginal return to savings for workers lacks the term deriving from profits, $\omega_a(0,a) = (1-\tau_a) + r(1-\tau_k)$.

it is welfare-maximizing since, as we will see next, it achieves some growth without excessively compromising equity. Output decreases as capital income tax rates increase due to misallocation. After reaching its maximum, output also starts to drop when capital income tax rates keep going down. This is a consequence of capital accumulation falling considerably.

The bottom panels look at the changes in inequality and welfare. The first two panels describe a very similar pattern for the gini indices of wealth and consumption. For lower levels of capital income taxation, entrepreneurs accumulate wealth with more ease and become closer to produce at their individually optimal size. At the same time the economy grows, resources become more concentrated in the hands of entrepreneurs both due to the strengthening of the self-financing channel and their firms obtaining larger profits. Consequently, the efficiency gains from taxing wealth instead of capital income are mitigated by the distributional consequences. The bottom right panel plots welfare in terms of consumption equivalence units. The optimal tax combination implies a modest decrease on the capital income tax rate (and an accompanying increase in the wealth tax rate) with respect to the status quo. The social planner seeks to maximize output, having to balance the effects of taxes on misallocation and aggregate wedges plus the savings incentives of households; that is, it looks for efficient capital usage and sufficient capital accumulation. At the same time, obtaining a larger economy pays off as long as the gains are properly redistributed, particularly at the bottom end of the consumption distribution. Opting for a smaller capital income tax rate would lead to a larger economy (unless it is too small, in which case the economy could also be smaller) where additional resources are mostly shared among the richest entrepreneurs. On the contrary, taxing capital income at higher rates would increase equity at the cost of downsizing the economy. This exercise calls for a positive combination of both fiscal instruments.

Last, we can think about a tax reform where the government sets the tax rate on capital income to zero and levies a flat-rate tax on wealth, while keeping taxes on labor income and consumption unchanged. This corresponds to the zero mark in all panels from Figure 1. As it is obvious from the welfare locus, the social planner would rather stay under the benchmark fiscal system than being in the steady-state proposed by this simple reform. Despite a 2% grow in output, the reform also implies an increase of consumption inequality (0.03 rise in its gini index), which sweeps away any efficiency improvements. This is in stark contrast with the results from Guvenen et al. (2023), which find the steady-state economy under this fiscal configuration to be welfare-improving.

Figure 1: Effect of tax reforms.




Notes: the left vertical line denotes the zero capital income tax rate, the middle vertical line denotes the capital income tax rate from the optimal combination, the right vertical line denotes the capital income tax rate from the baseline economy. Output is expressed in percentage deviations from the baseline economy; misallocation is expressed in percentage terms; welfare is expressed in consumption equivalence percentage units; the rest of panels are expressed in levels.

# 6 Conclusion and Discussion

Under heterogeneity in rates of returns, non-trivial differences between capital income and wealth taxation emerge. One extended approach to model this heterogeneity is through financial constraints on private firms. Under such environment, taxing wealth instead of capital income improves allocative efficiency by favouring wealth accumulation of productive entrepreneurs at the cost of rising wealth concentration at the top. This paper revisits the debate on whether to tax capital income or wealth in an economy where inefficiencies and heterogeneous returns also arise due to the presence of variable markups. To address the question, I build a model that reproduces notable US inequality features and market power dispersion.

I find that a social planner seeking welfare maximization and restricted to combinations of wealth and capital income taxation ensuring government revenue neutrality sets a positive rate on both taxes. Aggregate productivity gains from taxing wealth instead of capital income in an economy with financial frictions are smaller when accounting for variable markups. While misallocation coming from collateral is partially corrected by reducing capital income taxation, this is only one of the sources of inefficiencies in the economy. The same producers that overcome credit constraints generate inefficiencies by exploiting their market power in the form of larger markups and lower production quantities than absent variable markups. Although this second effect is less responsive to tax changes, it cannot be remediated with the same tools as the first one. Because entrepreneurs escape financial frictions and charge higher markups under low capital income taxation, they generate larger profits and resources become more concentrated. In addition, saving incentives ruled by general equilibrium effects and constrained entrepreneurs explain larger capital accumulation when capital income taxation is large. The social planner trades-off these opposing forces and finds slightly reducing taxation on capital income while subtly increasing wealth taxation optimal. This leads to a tax system much closer to the benchmark economy than to one where a considerably larger fraction of revenue is obtained via wealth taxation. Given the proximity between the optimal and original economies, welfare gains are modest although still positive.

The model in this paper focused on emphasizing the role of product market power in the debate about whether to tax capital income or wealth. As it is common in quantitative exercises, I abstracted from elements that may also be relevant for this exercise. To name a few, life-cycle dynamics, income dynamics, entrepreneurial risk or the presence of public firms; all of which could modify the results to some extent.

This analysis together with previous papers studying this question (Guvenen et al., 2023; Boar and Midrigan, 2023) highlights important lessons for policy work. Across these papers, discrepancies in results come from modelling choices that lead to different compositions of inefficiencies. An economy where entrepreneurial firms face strong financial limitations is likely to benefit from reducing capital income taxation and recovering revenue via other fiscal instruments. At the same time, models predict increased inequality under these fiscal policies. Thus, an inequality-averse social planner (or policy maker) should be cautious about the degree of inequality in the pre-reform economy and weight efficiency and equity consequences. This paper stressed another important ingredient: the plausible presence of additional inefficiencies and their interactions with tax changes. In economies where larger firms exert more market power over their competitors, overcoming financial frictions leads to higher markups, reducing overall efficiency gains and worsening equity motives. Even if this response is not large, it adds another force against a decrease in capital income taxation. Thus, the appropriate policy for an economy will depend on the diagnosed weaknesses.

# Bibliography


- Atkeson, A., & Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. American Economic Review, 98(5), 1998-2031.

- Bach, L., Calvet, L. E., & Sodini, P. (2018). Rich Pickings? Risk, Return, and Skill in Household Wealth. HEC Paris Research Paper No. FIN-2016-1126, Swedish House of Finance Research Paper, (16-03).

- Baqaee, D. R., & Farhi, E. (2020). Productivity and misallocation in general equilibrium. The Quarterly Journal of Economics, 135(1), 105-163.

- Boar, C., & Knowles, M. (2024). Optimal taxation of risky entrepreneurial capital. Journal of Public Economics, 234, 105100.

- Boar, C., & Midrigan, V. (2019). Markups and inequality (No. w25952). National Bureau of Economic Research.

- Boar, C., & Midrigan, V. (2023). Should We Tax Capital Income or Wealth?. American Economic Review: Insights, 5(2), 259-274.

- Brüggemann, B. (2021). Higher taxes at the top: The role of entrepreneurs. American Economic Journal: Macroeconomics, 13(3), 1-36.

- Buera, F. J., Kaboski, J. P., & Shin, Y. (2011). Finance and development: A tale of two sectors. American economic review, 101(5), 1964-2002.

- Cagetti, M., & De Nardi, M. (2009). Estate taxation, entrepreneurship, and wealth. American Economic Review, 99(1), 85-111.

- Carroll, C. D. (2006). The method of endogenous gridpoints for solving dynamic stochastic optimization problems. Economics letters, 91(3), 312-320.

- Conesa, J. C., Kitao, S., & Krueger, D. (2009). Taxing capital? Not a bad idea after all!. American Economic Review, 99(1), 25-48.

- De Loecker, J., Eeckhout, J., & Unger, G. (2020). The rise of market power and the macroeconomic implications. The Quarterly Journal of Economics, 135(2), 561-644.

- Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. The American economic review, 67(3), 297-308.

- Edmond, C., Midrigan, V., & Xu, D. Y. (2023). How costly are markups?. Journal of Political Economy, 131(7), 000-000.

- Fagereng, A., Guiso, L., Malacrino, D., & Pistaferri, L. (2020). Heterogeneity and persistence in returns to wealth. Econometrica, 88(1), 115-170.

- Guvenen, F., Kambourov, G., Kuruscu, B., Ocampo, S., & Chen, D. (2023). Use It or Lose It: Efficiency and Redistributional Effects of Wealth Taxation. The Quarterly Journal of Economics, 138(2), 835-894.

- Hsieh, C. T., & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. The Quarterly journal of economics, 124(4), 1403-1448.

- Imrohoroglu, A., Kumru, C., & Nakornthab, A. (2018). Revisiting tax on top income. manuscript, University of Southern California, 1, 3-2.

- Kimball, M. (1995). The quantitative analytics of the basic neomonetarist model. Journal of Money, Credit and Banking, 27(4), 1241–77.

- Klenow, P. J., & Willis, J. L. (2016). Real rigidities and nominal price changes. Economica, 83(331), 443-472.

- Kuhn, M., & Ríos-Rull, J. V. (2016). 2013 Update on the US earnings, income, and wealth distributional facts: A View from Macroeconomics. Federal Reserve Bank of Minneapolis Quarterly Review, 37(1), 2-73.

- Macnamara, P., Pidkuyko, M., & Rossi, R. (2022). Taxing Consumption in Unequal Economies (No. 2210). Economics, The University of Manchester.

- Midrigan, V., & Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. American economic review, 104(2), 422-458.

- Moll, B. (2014). Productivity losses from financial frictions: Can self-financing undo capital misallocation?. American Economic Review, 104(10), 3186-3221.

- OECD. (2018). The role and design of net wealth taxes in the OECD. OECD Tax Policy Studies, 26.

- Quadrini, V. (2000). Entrepreneurship, Saving, and Social Mobility. Review of Economic Dynamics, 3(1), 1-40.

# A Appendix

## A.1 Derivations

### A.1.1 Final Good Producer Demands

The final good producer solves

$$\max_{\{q_i\},L} Q^\alpha L^{1-\alpha} - \int_0^\psi p_i q_i di - WL \qquad \text{s.t.} \quad \int_0^\psi \Upsilon\Big(\frac{q_i}{Q}\Big) di = 1.$$

Define the function

$$F = \int_0^\psi \Upsilon\Big(\frac{q_i}{Q}\Big) dj - 1 = 0,$$

so that using the Implicit Function Theorem

$$\frac{dQ}{dq_i} = \Upsilon'\Big(\frac{q_i}{Q}\Big)\Bigg[\int_0^\psi \Upsilon'\Big(\frac{q_i}{Q}\Big)\frac{q_i}{Q} di\Bigg]^{-1} = \Upsilon'\Big(\frac{q_i}{Q}\Big) D.$$

It follows the first order conditions yield

$$p_i = \alpha\frac{Y}{Q}\Upsilon'\Big(\frac{q_i}{Q}\Big)D, \qquad W = (1-\alpha)\frac{Y}{L}.$$

### A.1.2 Entrepreneurs

Entrepreneurs solve

$$\begin{aligned} \pi(a,z) = \max_{q,p} & \left\{p(q;Q,D)q - \frac{r+\delta}{z}q\right\} \\ \text{s.t.} \quad & p = \alpha\frac{Y}{Q}\Upsilon'\Big(\frac{q}{Q}\Big)D = \Upsilon'\Big(\frac{q}{Q}\Big)X, \\ & q \leq \lambda z a. \end{aligned}$$

The FOC is

$$
\begin{aligned}
& p + q\frac{\partial p}{\partial q} = \frac{r+\delta}{z} + \Lambda \\
& \Rightarrow p + q\Upsilon''\Big(\frac{q}{Q}\Big)\frac{1}{Q}X = \frac{r+\delta}{z} + \Lambda \\
& \Rightarrow p + \underbrace{\Upsilon'\Big(\frac{q}{Q}\Big)X}_{p} \underbrace{\left[\frac{\frac{q}{Q}\Upsilon''\Big(\frac{q}{Q}\Big)}{\Upsilon'\Big(\frac{q}{Q}\Big)}\right]}_{-\left(\sigma(q/Q)^{-\varepsilon/\sigma}\right)^{-1}} = \frac{r+\delta}{z} + \Lambda \\
& \Rightarrow p = \frac{\sigma}{\sigma - (q/Q)^{\varepsilon/\sigma}}\Big(\frac{r+\delta}{z} + \Lambda\Big).
\end{aligned}
$$

### A.1.3 Aggregate Wedges

Combining the firm-level and aggregate conditions gives

$$
\frac{\mu(x)}{\mathcal{M}} = \frac{p(x)q}{P_Q Q}\frac{(1+\mathcal{T}^k)K}{(1+\tau^k)k}.
$$

Define the aggregate capital wedge as the capital-weighted average of individual capital wedges:

$$
1+\mathcal{T}^k = \int_{(s,z,a)} (1+\tau^k(z,a))\frac{k(z,a)}{K}d\Gamma(s,z,a).
$$

In the same fashion as Edmond, Midrigan and Xu (2015), rearranging and integrating on both sides, the aggregate markup can be expressed as a cost-weighted harmonic average of firm-level markups,

$$
\mathcal{M} = \left[\int_{s,z,a} \frac{1}{\mu(x(z,a))}\frac{p(x(z,a))q(z,a)}{P_Q Q}d\Gamma(s,z,a)\right]^{-1}.
$$

### A.1.4 Aggregate Productivity and Misallocation

Let $Z$ denote the aggregate productivity of the efficiency-adjusted capital sector, implicitly defined by an aggregate production function that relates the aggregate level of capital used in production $K$ to the total amount of efficiency-adjusted capital $Q$:

$$
Q = ZK.
$$

Combining the above expression with the firm-specific production function $q = zk$, $Z$ can be expressed as an harmonic mean of individual productivities weighted by their market power $x =$

$q/Q$,

$$\begin{aligned}
&\frac{q}{Q} = x = \frac{zk}{ZK} \\
&\Rightarrow \frac{x}{z} = \frac{1}{Z}\frac{k}{K} \\
&\Rightarrow Z = \Big[\int_{s,z,a} \frac{x(z,a)}{z} d\Gamma(s,z,a)\Big]^{-1}.
\end{aligned}$$

Taking the economy with CES aggregation (i.e. without variable markups) and without variable markups as the benchmark economy, following Hsieh and Klenow (2009) define the productivity of an economy without distortions as

$$Z^* = \Big[\int_{z,a} z^{\sigma-1}\Big]^{\frac{1}{\sigma-1}}.$$

It follows the degree of misallocation with respect to this benchmark economy can be computed as

$$\Omega = 1 - \frac{Z}{Z^*}.$$

#### A.1.5 Proof of Proposition 1

From the optimality condition of the entrepreneur (3), applying the chain rule it follows

$$\frac{\mathrm{d}\mu}{\mathrm{d}a} = \frac{\mathrm{d}\mu}{\mathrm{d}x} \cdot \frac{\mathrm{d}x}{\mathrm{d}\tau^k} \cdot \frac{\mathrm{d}\tau^k}{\mathrm{d}a}.$$

The derivative of markups with respect to relative size follows directly from the demand elasticity being decreasing in the relative size of the firm (see (1)), so $\mathrm{d}\mu/\mathrm{d}x > 0$.

The derivative of the relative firm size with respect to the capital wedge can be computed applying the Implicit Function Theorem. Let $F(x,\tau^k) = \Upsilon'(x)X - \mu(x)\frac{(1+\tau^k)(r+\delta)}{z} = 0$, so

$$\frac{\mathrm{d}x}{\mathrm{d}\tau^k} = \mu(x)\frac{r+\delta}{z} \cdot \Big[\Upsilon''(x)X - \frac{d\mu}{dx}\frac{(1+\tau^k)(r+\delta)}{z}\Big]^{-1} < 0.$$

Trivially, the capital wedge weakly decreases with wealth, $\mathrm{d}\tau^k/\mathrm{d}a \leq 0$. Thus, the derivative of markups with respect to wealth is weakly increasing. $\square$

D. Josep Pijoan Mas, Director del Programa de Doctorado en Economía y Gobierno que se imparte en el Centro de Estudios Monetarios y Financieros (CEMFI), Fundación del Banco de España, reconocida y clasificada por Orden Ministerial de 12 de abril de 1991 e inscrita en el Registro único de Fundaciones de competencia estatal con el número 182EDU, certifica:

Que D. Juan Segura, con DNI número 02311194Q, es alumno del Programa de Doctorado en Economía y Gobierno que se imparte en este Centro en régimen de dedicación exclusiva bajo mi supervisión desde julio de 2023. La fecha de conclusión del Programa de Doctorado se estima en julio de 2027.

Que el Programa de Doctorado en Economía del CEMFI es un título oficial de la Universidad Internacional Menéndez Pelayo (UIMP) de acuerdo con la regulación del Ministerio de Educación y Formación Profesional.

Y para que conste donde convenga al interesado, expido el presente certificado en Madrid, a la fecha de la firma digital.

Casado del Alisal 5 28014 Madrid Phone +34 91 429 0551 www.cemfi.es