ESCOLA BRASILEIRA DE ECONOMIA E FINANÇAS

Ensaios Econômicos

EPGE
Escola
Brasileira de
Economia e
Finanças

N° 839 ISSN 0104-8910

Reassessing the Efficiency-Equity Trade-off: Progressivity's Impact on Growth

Carlos E. da Costa, Artur Rodrigues

February de 2024

URL: https://hdl.handle.net/10438/5

Reassessing the Efficiency-Equity Trade-off:
Progressivity's Impact on Growth/ Carlos E. da Costa,
Artur Rodrigues - Rio de Janeiro: FGV EPGE, 2024
44p. - (Ensaios Econômicos; 839)

Inclui bibliografia.

CDD-330

Reassessing the Efficiency-Equity Trade-off: Progressivity's Impact on Growth*

Carlos E. da Costa — *EPGE FGV*

Artur Rodrigues — *EPGE FGV*

February 6, 2024

Abstract

We revisit the efficiency-equity trade-off of optimal tax theory by emphasizing the consequences of increased progressivity on *growth*. We use an endogenous growth framework that considers both the decision to become a researcher and the effort established entrepreneurs make to improve their products. We find that the optimal level or progressivity is lower than the current one but that welfare gains are moderate – $CEV \approx 0.5\%$. However, if one disregards the growth impact one would prescribe a substantially higher level of progressivity at significant welfare cost – $CEV \approx -9\%$. **JEL classification**: *O30;H21* **Keywords**: *Creative Destruction; Research Spillovers; Progressive Taxation*

OPTIMAL tax theory emphasizes the trade-off between equity and efficiency typically captured by the *level* of output that is sacrificed for a more equitable distribution of income. Yet, a potentially more consequential side-effect of the disincentives created by tax distortions, its impact on growth, remains understudied.

*Corresponding author: Artur Rodrigues artur.bfrodrigues@gmail.com. da Costa thanks CNPq project 304955/2022-1 for financial support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. We thank participants in the EPGE FGV Macro group and the 45th SBE meeting for their comments and suggestions. All errors are our responsibility.

We investigate the equity/growth trade-off in light of an endogenous growth model along the lines of Jones and Kim [2018]. A progressive income tax is used to redistribute income across individuals thus allowing risk to be shared across lucky and unlucky individuals with risky occupational choices and between these agents and those who opt for a safe occupation. Agents are heterogeneous in both their labor market productivity and their cost of engaging in research. This adds a distributive motive to the risk-sharing role of taxes. This added heterogeneity is key to capturing the entire distribution of income for the U.S., which is essential for our goal of quantifying the aforementioned equity/efficiency trade-off.

We focus on constant progressivity tax schedules – Benabou [2002] and Heathcote et al. [2017].[1] The policy experiment is to vary the progressivity parameter and adjust the level of taxes in such a way as to preserve the level of government consumption as a share of GDP fixed. More progressivity leads to less growth and inequality thus inducing an increasing relationship between the growth rate and the Gini coefficient. We find that the curve that represents this trade-off is strictly increasing and strictly convex; the more redistribution the greater the growth sacrifice that is required for further inequality reductions. Or, seen from the symmetric angle, the higher the growth rate the more one needs to sacrifice equity to further increase the growth rate.

Our central welfare assessment is done by identifying the Utilitarian optimal level of income tax progressivity. Intensifying progressivity in the tax system enhances income redistribution and fosters better risk-sharing among individuals. This comes with the traditional work disincentive emphasized in the macroeconomic literature. This is not all, however. Entrepreneurial research efforts contribute to technological advancements and stimulate the creative destruction process by increasing the gains from research by potential entrants. Progressive taxes reduce the payoff in the case of success and increase or decrease the payoff in the case of failure depending on where in the distribution of labor market productivity researchers are drawn from. We assess how these forces play out in an economy calibrated to reproduce the labor market productivity profile of researchers found in the data for the U.S. economy.

We find that the ideal level of progressivity is slightly lower than the current one. Yet, transitioning to the optimal tax system is expected to yield only moderate welfare gains.

[1] The schedule is automatically adjusted to keep the budget constraint satisfied. This precludes the income-weighted average marginal tax rate from mechanically changing with economic growth.

However, failing to account for the growth consequences of progressivity could lead to misguided policy prescriptions; ignoring the effects of progressivity on entrepreneurial research leads one to prescribe very high levels of progressivity, resulting in significant welfare losses – $CEV = -9.0\%$.

An important statistic for our quantitative exercises is the elasticity of productivity growth to the progressivity parameter. In our baseline exercise, we adopt a conservative (low) value, based on Akcigit et al.'s (2021) findings for the elasticity of innovation to taxes. Even lower elasticities can rationalize the current progressivity whereas a more central value would make the current welfare loss far from trivial, $CEV = -2.7\%$.

The rest of the paper is organized as follows. After a brief literature review, Section I displays the environment. In Section II, we present the definition of a balanced growth path equilibrium and its characterization. In Section III, we describe the calibration and present our main quantitative findings. Section IV concludes the paper. Longer proofs are collected in the appendix.

Lit. Review

The main question we are after in this paper relates to the consequences of taxation on economic growth. This is also what Jaimovich and Rebelo [2017] investigates. They consider a version of Romer's [1990] growth model in which researchers are heterogeneous in their innate talent. As in our case, taxes can affect growth rates. They show how non-linear effects arise from linear taxes due to the researcher's heterogeneity.

Similarly, Li and Sarte [2004] studies the growth impact of progressivity in an endogenous growth model using Rebelo's [1991] approach, which is substantially different from Jones and Kim's [2018] on which our approach is based.

The model's dynamics are mostly based on the process of creative destruction famously conceptualized by Joseph Schumpeter [1950]. Models of this kind gained prominence since the seminal works of Aghion and Howitt [1992] and Grossman and Helpman [1991] and have been used to explain economic phenomena ranging from rising top income inequality across time [Jones and Kim, 2018] to business dynamics [Klette and Kortum, 2004].[2] The most salient feature of such models is the process of innovation arising from the capitalistic competition in which entrepreneurs are constantly trying

[2] A comprehensive presentation of the relevant literature is found in Aghion et al. [2015].

to gain an innovative edge for themselves to the detriment of their competitors. They do so to attain a position of monopoly power that could compensate for the uncertainty involved in their enterprises.[3] Thus, both entrant and incumbent entrepreneurs devote themselves to innovation. This is the driving force of economic growth in our model.

Acemoglu and Cao [2015] develops a model with a complex interplay between entrant and incumbent innovation; the former, is mainly disruptive, and the latter, is incremental. Acemoglu and Cao [2015] creates different growth decompositions and replicates Zipf's law for firm sizes. Garcia-Macia et al. [2019] estimates the contribution of new varieties of products, creative destruction, and firms' own-variety improvements using data on establishments and job creation and destruction aiming at giving an empirical account of these different sources of growth. They find that creative destruction could directly account for about one-fourth of economic growth.

Our model introduces progressive taxation in an environment built on Jones and Kim' (**?**). For our purposes producing an accurate income distribution is paramount so we add worker heterogeneity to further improve its empirical fit in the cross-sectional dimension. Indeed, Jones and Kim's model is well suited to account for the top of the income distribution, but not for the bottom. Because a substantive fraction of the benefits associated with a progressive system accrues to those at the lower end of the distribution, we can think of our extension as focused on getting the benefits right.

Closest to our work is Jones [2022] which investigates how innovation as a nonrival good shapes the analysis of top-income taxation and shows that revenue or welfare gains from (semi-endogenous) growth can severely limit top tax rates. As Jaimovich and Rebelo [2017], we consider an economy that produces endogenous long-run growth, even with a constant population. We also give the model a more complete mechanism of responses by considering both extensive and intensive margins — while those two papers consider either one or the other —, and include risk as a defining feature of the pursuit for innovation.

In a very early version of this paper, Rodrigues [2021] goes after some of the same questions we explore, but in a world without worker heterogeneity. He finds substantially lower progressivity to be optimal, which emphasizes the importance of taking

[3] In the words of Schumpeter himself (1950, p. 89): "enterprise would in most cases be impossible if it were not known from the outset that exceptionally favorable situations are likely to arise which if exploited by price, quality and quantity manipulation will produce profits adequate to tide over exceptionally unfavorable situations provided these are similarly managed".

distributive motives into account.

I The Model

The core of our model is a variation on Jones and Kim's [2018] Schumpeterian economy.[4] We enrich their framework by assuming agent heterogeneity to add a distributive motive to the use of progressive income taxes. Time is continuous denoted by t. Whenever convenient, we omit the time subscript from the variables.

I.1 Demographics and preferences

The economy is inhabited by a continuum of infinitely-lived utility-maximizing individuals of fixed size, N. We, therefore, assume no population growth. Each individual is characterized by a pair of parameters (ν, κ), where ν is the individual's labor productivity and κ is her cost of entering research.

Types are assigned to individuals through i.i.d. draws from the distributions,

$$\log \nu \sim \mathcal{N}(0, \sigma_\nu^2) \quad \text{c.d.f. } F_\nu(\nu)$$
$$\kappa \sim \text{Exp}(\psi) \quad \text{c.d.f. } H(\kappa)$$

We assume that, before entering the labor market, ν-workers face a lottery that, with a $b(\nu)$ probability of success, determines whether they will be able to choose to become researchers,

$$\chi(\nu) \in \{0,1\}, \quad \Pr(\chi(\nu) = 1) = b(\nu)$$
$$b(\nu) \in [0,1], \quad b'(\nu) \geq 0.$$

Under this assumption, a fraction, $1 - b(\nu)$, of each type, ν, faces no occupational choice. This two-stage determination of each agent's prospect allows us to keep the model tractable by making the conditional (on being a potential researcher) distribution of researchers independent of ability types while accommodating the empirical ratio of

[4]Jones and Kim [2018] in turn builds on an older tradition of modeling and understanding entrepreneurship and innovation – e.g., Aghion and Howitt [1992], Grossman and Helpman [1991], Schumpeter [1950].

workers and researchers for each type. Note that our assumption does not trivialize the distribution of researchers, since a positive mass of agents still decide on their occupation.

Individual preferences are defined over random paths of consumption and effort. In t, the associated expected utility is,

$$\mathbb{E}_t \int_t^{\infty} e^{-\rho(s-t)} \left(\log c_s - \frac{\ell_s^{1+\eta}}{1+\eta} \right) \mathrm{d}s, \tag{1}$$

in which $c_t \geq 0$ and $\ell_t \geq 0$ denote consumption and effort, respectively, and ρ is the discount rate. The parameter η corresponds to the inverse of the Frisch elasticity of effort. Effort ℓ by a (ν, κ)-agent generates $l = \nu\ell$ efficiency units of labor as a worker.

I.2 Production technology

There are two sectors in the economy: one produces final consumption goods from intermediate goods, and the other produces intermediate goods from labor. The final goods sector consists of a price-taking representative firm that combines a unit-measure continuum of varieties of intermediate goods according to the CES technology,

$$Y(\mathbf{q}) = \left(\int_0^1 q_i^{\theta} \, \mathrm{d}i \right)^{\frac{1}{\theta}}, \quad 0 < \theta < 1, \tag{2}$$

where $\mathbf{q} \equiv \{q_i\}$. The index $i \in [0, 1]$ specifies the goods' variety, and θ governs the degree of substitution between them.

These intermediate goods are in turn produced by entrepreneurs. The measure of goods' varieties is fixed and each entrepreneur owns the exclusive right to produce exactly one of them at a time. Their variety q_i is produced proportionally to effective labor hired, L_i,

$$q_i(L_i) = A_{it} L_i. \tag{3}$$

The time subscript in the factor productivity, A_{it}, anticipates its central role in the model dynamics. As we shall see this is ultimately related not only to inequality among entrepreneurs but also to the diffusion of ideas and their non-rival nature which, in the spirit of Romer [1990], enables growth. The determinants of A_{it} are explained next.

I.3 Entrepreneurship

Those engaged in entrepreneurship, broadly speaking, are subdivided into two categories representing their current state: *established entrepreneurs*, who own the right to produce a variety and do so under the conditions previously described; and *researchers*, who are trying to come up with new and better marketable ideas and displace the established entrepreneurs who occupy limited market space. We discuss each one in sequence. First, the established entrepreneurs, which we refer to simply as entrepreneurs.

An entrepreneur's productivity A_{it} is given by

$$A_{it} = \bar{A}_t x_i^{\phi}, \quad \phi > 0, \tag{4}$$

where $\bar{A}_t$ measures the aggregate technological level and x_i, the idiosyncratic productivity. We explain x_i in this section.

The relative position of the entrepreneur in the market, which she hopes to improve through effort, is measured by her idiosyncratic productivity denoted by x. For a given i, x is governed by the law of motion,

$$dx_{it} = \mu(\ell_{it})x_{it}dt + \sigma x_{it}dB_{it}, \quad dB_{it} \sim \mathcal{N}(0, dt), \tag{5}$$

a geometric Brownian motion with mean growth rate $\mu(\ell_{it}) \geq 0$ and percentage volatility $\sigma > 0$. Since μ is an increasing function of ℓ, the growth rate of x can thus be influenced (in expected value) by the effort exerted by the entrepreneur. Incumbent research, therefore, takes the form of an improvement effort.

Every new business starts with a normalized base productivity $x^0 = 1$. If left unchecked, x would grow without bounds unless a resetting mechanism acts as a stabilizing force. With suitably chosen resetting mechanisms x converges to well-known distributions.[5] Entrepreneurs can retire, die, or simply be competitively displaced by new entrepreneurs. Above all, in a free market economy the stronger the competition, the harder it is to maintain one's business profitable and the more likely it is for old entrepreneurs to go out of business. Here, we model exit as the first occurrence between two independent Poisson processes, faced equally by all entrepreneurs. One has an exogenous arrival rate $\bar{\delta}$ while the other has a rate δ^{cd} increasing in "outside pressure",

[5]E.g., Gabaix et al. [2016].

more precisely defined in the next section. Let us call the former process *natural retirement* and the latter *creative destruction* and denote the total resulting exit rate faced by an entrepreneur by[6]

$$\delta_t \equiv \bar{\delta} + \delta_t^{cd}. \tag{6}$$

Once the entrepreneur of a given variety, i, exits, the production of i is taken over by a new entrepreneur. If that happens through the natural retirement process, the variety's x_{it} accumulated up to that point is destroyed and set back to its base level x^0 providing the necessary resetting mechanism. We have, thus, fully described the "life-cycle" of an entrepreneur.

I.4 Research

Established entrepreneurs were researchers at some time in the past. Since the economy is served by a fixed unit mass of intermediate varieties, those involved in research are actively trying to come up with a better version of an existing variety to take the place of an incumbent entrepreneur. This research is undirected, that is, not tied to a particular variety, and success comes randomly at a Poisson rate $\bar{\lambda}$ for any given researcher.

If successful, the researcher gains exclusive production rights over a better-quality version of a randomly defined existing variety, rendering its former version obsolete. The entrepreneur who produced that old version must now become a researcher again to regain the position. This corresponds to the previously mentioned process of creative destruction. So, equating entrepreneurial entry and exit flows, and letting R_t denote the measure of researchers, we get

$$\delta_t^{cd} = \bar{\lambda} R_t.$$

Entrepreneurs are, therefore, subject to the pressure of competition from the outside: the more outside research and innovation there is, the faster the rate at which they go out of business. Importantly, it is not within the powers of an entrepreneur to change the probability with which he or she is displaced.

A second mechanism through which researchers replace entrepreneurs can be described as follows. With identical probability, $\bar{\delta}/R$, across researchers, they are selected

[6] Note that natural retirement refers to the product or service that the entrepreneur is offering. The agent himself returns to the drawing board to try to have an idea that will allow him or her to return to being an entrepreneur.

to take over varieties left by entrepreneurs through the natural retirement process. Although the switch in positions has the same consequence from a private perspective, this form of replacement does not generate innovation like creative destruction does.

Taking these two processes, creative destruction and replacement of vacant variety, into account, the total rate at which researchers individually become entrepreneurs is given by

$$\lambda_t \equiv \bar{\delta}/R_t + \bar{\lambda}. \tag{7}$$

This is the counterpart to the definition in equation (6) for entrepreneurs.

The cost of researching is that agents must sacrifice their labor market production. In particular, we assume that a (ν, κ)-agent who is engaged in research generates only $\bar{l} = \xi\nu\ell$ efficiency units of labor for an effort ℓ, with $0 < \xi < 1$ common to all. As previously mentioned, it costs, in addition, κ units of utility to engage in research. Note that in this latter respect, types differ.

I.5 Innovation

Innovation is the sole responsible for long-term growth since it can indefinitely expand the stock of ideas and technology which determine the total factor productivity of the economy. To differentiate the economy's technological level from entrepreneurs' idiosyncratic productivity x_i, we use terms such as *quality* and *technology* to refer to them. However, they play the same role of a labor-augmenting factor in production.

Innovation occurs through two processes already described: creative destruction and incumbent research and is spread in the economy as a side effect. When creative destruction occurs, i.e., when the entrant entrepreneur comes up with a new idea, the technological level of a given variety i is raised, increasing A_{it} by a factor of $\gamma > 1$, the step size of innovation. We assume for simplicity that technological diffusion, a positive spillover effect of innovation, is instant and universal so that *all* varieties have their technological level raised by this same factor.

Incumbent research done by entrepreneurs, i.e. x-increasing effort, also contributes to innovation and also has a spillover effect to varieties other than their own. Likewise, it generates an increase by factor γ to all varieties' quality but weighted in proportion to their own-productivity growth rate $\mu(\ell_{it})$.

Let n_t denote the cumulative stock of innovation steps at time t and assume $n_0 = 0$.

We can define $\bar{A}_t$ in equation (4) as $\bar{A}_t \equiv \gamma^{n_t}$ and get the final form of A_{it},

$$A_{it} = \gamma^{n_t} x_i^{\phi} \tag{8}$$

The total contribution of researchers and entrepreneurs towards innovation at any given time is then,

$$\dot{n}_t = \delta^{cd}(R_t) + \iota \mu_t, \tag{9}$$

where $\iota > 0$ and $\mu_t \equiv \int \mu(\ell_{it})\,\mathrm{d}i$.

Since we have a uniform effect of innovation on all varieties, we get to simplify the analysis by having a single n_t that tracks the common progress of technology. This eliminates the need for integration across varieties on different levels of quality.

I.6 Market arrangements

The final good is the numeraire of the economy and the firms that produce it are perfectly competitive. Entrepreneurs, in contrast, operate in monopolistic competition, setting prices p_i for their products and hiring labor from workers in a competitive market that pays w per unit of effective labor.

There are no traded assets in the economy.

I.7 Government

The government redistributes income through taxes and transfers. We adopt the constant progressivity tax schedule of Benabou [2002] and others, in which the disposable income of household h is defined by

$$\hat{y}_{h,t} \equiv y_{h,t}^{1-\tau} \tilde{y}_t^{\tau}, \tag{10}$$

where $y_{h,t}$ denotes pre-tax income and the break-even income level $\tilde{y}_t$ is determined as a result of the government balancing its budget constraint. The implied tax function for income $y_{h,t}$ is, therefore,

$$T(y_{h,t}) \equiv y_{h,t} - y_{h,t}^{1-\tau} \tilde{y}_t^{\tau}, \tag{11}$$

and the government must raise a fraction $\zeta \geq 0$ of the economy's output in tax revenue so that it must satisfy at every t the following budget constraint:

$$\int_h T(y_{h,t})\,\mathrm{d}h = \zeta \int_h y_{h,t}\,\mathrm{d}h. \tag{12}$$

The policy parameter, $\tau \leq 1$, governs the progressivity of the schedule, which can be seen in the elasticity of disposable to pre-tax income, $(1-\tau)$. If $\tau > 0$, then the schedule is progressive and both marginal and average tax rates are increasing in pre-tax income. Otherwise, it is either regressive ($\tau < 0$) and the opposite occurs, or when $\tau \to 0$, taxes approach a flat rate schedule. Note also that, from a macro perspective, τ is equivalent to the income-weighted average marginal tax rate:

$$\int_h T'(y_{h,t}) \left(\frac{y_{h,t}}{Y_t}\right) \mathrm{d}h = \tau.$$

Finally, it is worth noting that the term $\tilde{y}_t^\tau$ resulting from balancing the budget ensures tax rates are always relative to average income so that the schedule automatically adjusts for the general income growth.

I.8 The household's problem

The household initially chooses the occupation: either worker or researcher. As mentioned, individuals are endowed only with their capacity to exert effort. Moreover, there are no financial assets available for them to smooth consumption. A ν-worker can convert a unit of effort into ν units of effective labor. These are, then, used in the production of intermediate goods and are paid a wage, w.

Researchers engage in the process of research already described, but must also supply work in the market to finance their consumption. They have, however, an opportunity cost of not working full-time, so that effort is converted at a rate $\xi\nu < \nu$ into effective labor.

Denote by

$$V_t(\nu,\kappa) \equiv \max\{V_t^W(\nu), V_t^R(\nu) - \kappa\} \tag{13}$$

the (ν,κ)-non-entrepreneur household value in period t. $V_t^W(\nu)$ and $V_t^R(\nu)$ are, respectively, the value of being a worker and of being a researcher, net of κ. Furthermore, let

$V_t^E(x)$ be the value of an entrepreneur with productivity x; then we can represent the problem recursively through the Bellman equations,

$$\rho V_t^W(\nu) = \max_{\ell_t} \; \log c_t^W(\ell_t, \nu) - \frac{\ell_t^{1+\eta}}{1+\eta} + \frac{dV_t^W(\nu)}{dt}, \tag{14}$$

$$\begin{aligned} \rho V_t^R(\nu) &= \max_{\ell_t} \; \log c_t^R(\ell_t, \nu) - \frac{\ell_t^{1+\eta}}{1+\eta} + \frac{dV_t^R(\nu)}{dt} \\ &\quad + \bar{\lambda}\left(\mathbb{E}_x[V_t^E(x,\nu)] - V_t^R(\nu)\right) + \bar{\delta}/R_t\left(V_t^{E_0}(\nu) - V_t^R(\nu)\right). \end{aligned} \tag{15}$$

where consumption is equal to disposable income, i.e., $c_t^W(\ell_t, \nu) = (w_t \nu \ell_t)^{1-\tau}\, \tilde{y}_t^\tau$ and $c_t^R(\ell_t, \nu) = (w_t \xi \nu \ell_t)^{1-\tau}\, \tilde{y}_t^\tau$. We use the shorthand $V_t^{E_0}(\nu) \equiv V_t^E(x^0, \nu)$. The last two terms in equation (15) account for the researchers' expected value change due to their becoming entrepreneurs in both possible ways.

With log-utility, the optimal choice of effort for both occupations is invariant to productivity, time, and wage rate: $\ell^W = \ell^R = (1-\tau)^{\frac{1}{1+\eta}}$.

The entrepreneur's value function is somewhat more involved since it has a diffusion process for its state variable. That is,

$$\rho V_t^E(x_t, \nu) = \max_{\ell_t} \left\{ \log c_t^E(x_t) - \frac{\ell_t^{1+\eta}}{1+\eta} + \frac{\mathbb{E}_t[dV_t^E(x_t,\nu)]}{dt} + \delta_t \left[V_t^R(\nu) - V_t^E(x_t,\nu)\right] \right\}. \tag{16}$$

Given the law of movement of x_t in equation (5), the expected rate of change term is established using Ito's lemma,

$$\frac{\mathbb{E}_t[dV_t^E(x_t,\nu)]}{dt} \equiv \mu(\ell_t) x_t \frac{\partial V_t^E(x_t,\nu)}{\partial x} + \frac{\sigma^2}{2} x_t^2 \frac{\partial^2 V_t^E(x_t,\nu)}{\partial x^2} + \frac{\partial V_t^E(x_t,\nu)}{\partial t}. \tag{17}$$

The entrepreneur's consumption c_t^E does not depend on ℓ_t. Instead, it is a function of productivity x_t which ultimately determines his or her income and is influenced by the agent's history of efforts. That is, the way effort creates value for the entrepreneur is by influencing their projected value growth through x_t, as made evident by the first term of equation (17). Therefore, ℓ_t plays a fundamentally different role for entrepreneurs than it does for workers and researchers, a role that more closely resembles "investing into one's business" than anything else.

The derivation of the optimal decision rule depends on the effort-converting technology $\mu(\ell_t)$ as well as the relationship between x_t and income, which still needs explaining.[7]

II Equilibrium

Let W_t be the measure of agents who either have no option but to be workers or choose to be, and R_t be the measure of agents who choose to engage in entrepreneurial activities but who have not yet become entrepreneurs. Then, we define a ***balanced growth path equilibrium – BGP*** as follows.

BGP Equilibrium Given a government policy, τ, a balanced growth path equilibrium consists of a measure of workers and researchers $\{W_t(\nu), R_t(\nu)\}$ of each type; choices of effort $\{\ell_t^W(\nu), \ell_t^R(\nu), \ell_t^E(x)\}$; intermediate input lists $\{\mathbf{q}_t\}$; tax levels $\{\tilde{y}_t\}$; a distribution $f(x)$; growth rate g; and prices $\{w_t, \{p_{it}\}\}$ which satisfy the following conditions for every t:

1. (ν, κ)-agents solve their recursive problems: $V_t(\nu, \kappa)$ for non-entrepreneurs, and; $V_t^E(x, \nu)$ for entrepreneurs, defining the associated $\ell_t^W, \ell_t^R, \ell_t^E(x)$ decision rules.[8]

2. The final goods firm maximizes profit by choosing inputs $\mathbf{q}_t = \{q_{it}\}$ given prices $\{p_{it}\}$; the intermediate goods firms maximize profits given wage rate w_t by setting prices $\{p_{it}\}$ under monopolistic competition.

3. Prices $\{p_{it}\}$ clear the intermediate goods market; wage rate w_t clears the labor market;

[7]While the utility of an entrepreneur with a higher ν is higher than that of an agent with a lower ν through V_t^R this has no bearing on the entrepreneur's choice of effort ℓ, hence on $\mathbb{E}_t[dV_t^E(x_t)/dt]$. The probability δ_t of being replaced is independent of x_t, hence, of ℓ. As a consequence, the additional expected utility that an entrepreneur gets from having a higher ν cannot be altered by his or her choices.

[8]Using the convention that an agent who did not get the chance of becoming a researcher, drew a value $\kappa = \infty$, the definition also applies to these agents.

4. W_t, R_t satisfy the population constraint,

$$W_t + R_t + 1 = N; \tag{18}$$

where W and R are endogenously defined as a result of (13);

5. The government chooses $\tilde{y}_t$ to balance its budget as in equation (12);

6. $f(x)$ is the stationary distribution of x, which satisfies the proper Kolmogorov forward equation, and;

7. Output grows at the constant rate of g.

To characterize the equilibrium, we now explain how each of the aggregate variables for which we will impose market clearing is determined.

II.1 Labor Supply

Agents in our economy make two types of choices: whether to remain workers or become entrepreneurs/researchers (extensive margin) and how much effort to make conditional on those choices (intensive margin). Let us start with the former.

II.1.1 The Extensive Margin

The net value of entering research for type (κ, ν) is

$$\mathcal{D}(\kappa, \nu) \equiv -\kappa/\rho + V^R(\nu) - V^W(\nu).$$

This defines thresholds, $\kappa^*(\nu)$, through $\mathcal{D}(\kappa^*(\nu), \nu) = 0$, that we later show to be given by

$$\kappa^*(\nu) = -\underbrace{\frac{(1-\tau)\lambda}{\rho + \lambda + \delta}}_{a} \log \nu + \underbrace{\rho[V^R(1) - V^W(1)]}_{C} = -a \log \nu + C,$$

in equilibrium.

We use this expression to define the share of each type, ν, with the opportunity to become a researcher that does become one, $r^*(\nu) \equiv \Pr\left(\mathcal{D}(\kappa, \nu) > 0 \mid \nu,\ \chi(\nu) = 1\right)$. In this case, $r(\nu) = b(\nu)r^*(\nu)$ is the share of ν-types that become researchers, and $1-r(\nu)$, is the share that are workers. Following from the definitions and the distribution of κ,

$$r^*(\nu) = H(\kappa^*(\nu)) = \begin{cases} 1 - e^{-\psi\kappa^*(\nu)} & k^*(\nu) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} 1 - e^{-\psi C}\nu^{\psi a} & 0 \leq \nu \leq e^{C/a} \\ 0 & \text{otherwise} \end{cases}$$

Those who engage in research eventually stop contributing with their effort to the production of varieties as they become entrepreneurs. To characterize the aggregate supply of efficiency labor units, we return to the population constraint, $N = W + R + 1$. Those who have no option but to remain workers, comprise a share $1 - b(\nu)$ for each ν. Those who can become researchers but choose not to, comprise a share $b(\nu)(1 - r^*(\nu))$ for each ν. Adding the two groups we obtain the number of workers in the population,

$$W = N \int \left[1 - b(\nu) + b(\nu)(1 - r^*(\nu))\right] f_\nu(\nu)\mathrm{d}\nu$$
$$= \int \underbrace{\left[N(1 - r(\nu))f_\nu(\nu)\right]}_{\equiv W(\nu)} \mathrm{d}\nu = \int W(\nu)\mathrm{d}\nu. \quad (19)$$

Among those who are engaged with entrepreneurship at any given time, one must distinguish those in the research phase, who still contribute to the labor force, from established entrepreneurs, who do not:

$$R + 1 = N \int \left[b(\nu)r^*(\nu)f_\nu(\nu)\right] \mathrm{d}\nu$$
$$\therefore \quad R = \int \underbrace{\left[N \frac{R}{R+1} r(\nu)f_\nu(\nu)\right]}_{\equiv R(\nu)} \mathrm{d}\nu = \int R(\nu)\mathrm{d}\nu. \quad (20)$$

$R(\nu)$, the measure of researchers at each ν, is defined this way because, in equilibrium, the ratio of researchers to entrepreneurs must be the same across ν, since all face the same rate of entry and exit.

II.1.2 The Intensive Margin

If an agent opts to be or is constrained to be a worker, or if the agent is in the research phase of life as an entrepreneur, the effort choice is essentially static, and given by

$$\max_{\ell}(1-\tau)\ln\ell - \ell^{1+\eta}/(1+\eta). \tag{21}$$

The choice depends neither on the productivity ν nor on whether the agent is a full-time worker or whether some of what he/she produces is spent on research activities, thus leading to $\ell^W(\nu) = \ell^R(\nu) = \ell^W$, where ℓ^W is the solution of (21).

Because established entrepreneurs spend no time in production, knowing ℓ^W, is sufficient to characterize the aggregate supply of efficient labor,

$$L = \ell^W \left(\int \nu W(\nu)\mathrm{d}\nu + \xi \int \nu R(\nu)\mathrm{d}\nu \right) \equiv \ell^W \left(\bar{\nu}^W + \xi \bar{\nu}^R \right),$$

noting that $\bar{\nu}$ denotes aggregate ν.

The intensive margin choice for entrepreneurs is more involved. First, we cannot ignore its dynamic nature: effort affects the expected growth of x, thus all the future consumption path. Second, because the whole path is affected, the assumption of $\ln$ preferences does not eliminate the impact of equilibrium aggregate variables on choices.

Hence, to characterize these choices, in all that follows we assume that

$$\mu(\ell) = \beta\ell^{1-\alpha}, \quad \beta > 0,\ \alpha < 1. \tag{22}$$

Moreover, since we are interested in equilibria with interior solutions for constant (W, R) we assume for now that this is the case to derive Proposition 1, below. Later, we show that consistency with a balanced growth path requires the constancy of (W, R).

Proposition 1. *Given the equilibrium measure of researchers, R, the optimal entrepreneurial effort on the balanced growth path is*

$$\ell^E = \left(\frac{\beta(1-\alpha)(1-\tau)}{\rho + \delta(R)} \right)^{\frac{1}{\eta+\alpha}}. \tag{23}$$

Two observations about proposition 1 are in order. First, the only endogenous variable that affects entrepreneurial effort is the number of researchers, through the creative destruction rate. An immediate consequence is that a constant R implies a constant ℓ^E. Second, the elasticity of entrepreneurial effort to tax progressivity, $-\tau/[(\eta+\alpha)(1-\tau)]$, is higher in absolute value than that of workers and researchers, $-\tau/[(\eta+1)(1-\tau)]$.

II.2 Output, wages, and profits.

Proposition 2. *Let $X_t \equiv \int x_{it}\,\mathrm{d}i$. Given the market arrangements previously described and assuming technology parameters are such that $\phi = (1-\theta)/\theta$, total output in period t is given by*

$$Y_t = \gamma^{n_t} X_t^{\phi} L_t, \tag{24}$$

where $L_t \equiv \int L_{it}\,\mathrm{d}i$, the wage rate by

$$w_t = \theta\gamma^{n_t} X_t^{\phi}, \tag{25}$$

and profits for the entrepreneur i by

$$\pi_{it} = (1-\theta)\gamma^{n_t} X_t^{\phi-1} L_t x_{it}. \tag{26}$$

The assumption about ϕ and θ guarantees that profits are linear in x_{it}, which implies that x_{it} and income share the same distribution and greatly simplifies the algebra. It is by no means essential for our results. Now, using the linearity, we can write the entrepreneur's pre-tax income as $y_t^E(x_{it}) = m_t x_{it}$ with m_t defined to conform to (26):

$$m_t \equiv (1-\theta)\gamma^{n_t} X_t^{\phi-1} L_t. \tag{27}$$

Note also that, given X^{ϕ}, wages are fully determined by the demand side. While wages do not directly affect the intensive margin of labor supply, it is important for the extensive margin choices.

II.3 The Stationary Productivity Distribution

Let $f(x,t)$ be the probability density function of x at time t and take $f(x,0)$ as given. Assume a fixed $\mu(\ell_{it}) = \mu^*$ common to all entrepreneurs, which we prove to be the case in equilibrium. Furthermore, we assume for simplicity that the base productivity $x^0 = 1$ is also the minimum productivity possible, i.e. there is a "reflecting barrier" at x^0 which precludes x_t from getting lower than it.[9] Under these conditions, outside the point of re-injection, x^0, the distribution $f(x,t)$ satisfies the following Kolmogorov forward equation,

$$\frac{\partial f(x,t)}{\partial t} = -\bar{\delta} f(x,t) - \frac{\partial}{\partial x}[\mu^* x f(x,t)] + \frac{1}{2} \cdot \frac{\partial^2}{\partial x^2}[\sigma^2 x^2 f(x,t)]. \tag{28}$$

If a stationary distribution $f(x) \equiv \lim_{t\to\infty} f(x,t)$ exists, then it must satisfy

$$0 = -\bar{\delta} f(x) - \frac{\partial}{\partial x}[\mu^* x f(x)] + \frac{1}{2} \cdot \frac{\partial^2}{\partial x^2}[\sigma^2 x^2 f(x)], \tag{29}$$

whence we derive Proposition 3, below.

Proposition 3. *The stationary distribution of x which satisfies (29) is the Pareto distribution,*

$$f(x) = \begin{cases} z^* x^{-z^*-1} & \text{for } x > 1 \\ 0 & \text{for } x < 1 \end{cases}, \tag{30}$$

where

$$z^* = \frac{-\tilde{\mu}^* + \sqrt{(\tilde{\mu}^*)^2 + 2\sigma^2 \bar{\delta}}}{\sigma^2}, \tag{31}$$

and $\tilde{\mu}^* \equiv \mu^* - \sigma^2/2$.

This establishes both the stationary distribution of entrepreneurs' income and the aggregate (or mean) productivity X_t, provided: *i)* the shape parameter z^* is larger than one, so that $f(x)$ has a finite mean, and; *ii)* $\mu(\ell_{it})$ is constant in equilibrium.

Simple inspection allows us to verify i). As for ii), assume for the moment that $R_t = R$, constant. Then the value functions are characterized in the following proposition.

Proposition 4. *The value functions of workers, researchers, and entrepreneurs in a BGP*

[9] More precisely, it means $x_{t+dt} = \max\{x^0, x_t + dx_t\}$ for small dt. It is possible to relax this assumption — we get a stationary double Pareto distribution [Reed, 2001] —, but the algebra gets more unwieldy for little to no added insight.

equilibrium are respectively given by

$$\rho V_t^W(\nu) = gt + \frac{g}{\rho} + \tau \log \tilde{y} + v^W(\nu) \tag{32}$$

$$\rho V_t^R(\nu) = gt + \frac{g}{\rho} + \tau \log \tilde{y} + \frac{(\rho+\delta)v^R(\nu) + \lambda v^E}{\rho+\lambda+\delta} \tag{33}$$

$$\rho V_t^E(\nu, x) = gt + \frac{g}{\rho} + \tau \log \tilde{y} + \frac{(\rho+\lambda)v^E + \delta v^R(\nu)}{\rho+\lambda+\delta} + \frac{\rho}{\rho+\delta}\log x \tag{34}$$

where $v^W(\nu), v^R(\nu), v^E$ are related to the value flows that are particular to the time spent in each occupation:

$$v^W(\nu) = (1-\tau)\log(w_0 \nu \ell^W) - \frac{(\ell^W)^{1-\eta}}{1-\eta}, \tag{35}$$

$$v^R(\nu) = (1-\tau)\left[\log(\xi w_0 \nu \ell^R) + \frac{\bar{\lambda}/z}{\rho+\delta}\right] - \frac{(\ell^R)^{1+\eta}}{1+\eta}, \tag{36}$$

$$v^E = (1-\tau)\left[\log(m_0) + \frac{\tilde{\mu}}{\rho+\delta}\right] - \frac{(\ell^E)^{1+\eta}}{1+\eta}.. \tag{37}$$

Now, if R_t is constant, then, ℓ_t^E is constant due to Proposition 1 and growth is constant due to (9). Direct inspection of Equations (32) – (37) allows us to see that $v^W(\nu) - V_t^R(\nu)$ is constant for all ν. From (13), so is R_t, justifying our guess.

Before we move to the quantitative analysis, a remark on the "weights" of v^E and v^R in equations (33) and (34) is due. In a continuous-time Markov chain with two states $\{1, 2\}$ and rate matrix

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \delta & -\delta \end{bmatrix} \tag{38}$$

the probability matrix $P(t)$ with entries $p_{ij} = Prob(\chi_t = j | \chi_0 = i)$, where χ_t denotes the state at t, is given by

$$P(t) = \begin{bmatrix} \frac{\delta}{\delta+\lambda} + \frac{\lambda}{\delta+\lambda}e^{-(\lambda+\delta)t} & \frac{\lambda}{\delta+\lambda} - \frac{\lambda}{\delta+\lambda}e^{-(\lambda+\delta)t} \\ \frac{\delta}{\delta+\lambda} - \frac{\delta}{\delta+\lambda}e^{-(\lambda+\delta)t} & \frac{\lambda}{\delta+\lambda} + \frac{\delta}{\delta+\lambda}e^{-(\lambda+\delta)t} \end{bmatrix}. \tag{39}$$

The coefficients of v^R and v^E, fundamentally acting as weights, in equations (64) and (65) correspond to the present value (times ρ), discounted at ρ, of the probabilities in $P(t)$, that is,

$$\rho \int_0^\infty e^{-\rho t} P(t) \mathrm{d}t = \begin{bmatrix} \frac{\rho+\delta}{\rho+\lambda+\delta} & \frac{\lambda}{\rho+\lambda+\delta} \\ \frac{\delta}{\rho+\lambda+\delta} & \frac{\rho+\lambda}{\rho+\lambda+\delta} \end{bmatrix}. \tag{40}$$

Thus, we have, for example, a researcher weighting the value v^R according to the present value of the probability he remains in research through time.

III Quantitative Findings

III.1 Parametrization

The model's parameters are either calibrated using empirical targets or taken from estimates and conventional values in the literature, assuming throughout that a unit of time t corresponds to a year. All moments are calculated from the BGP equilibrium variables. We also assume hereafter the functional form

$$b(\nu) = 1 - e^{-\bar{b}\nu}, \quad \bar{b} > 0. \tag{41}$$

The full baseline parametrization is summarized in table 1.

Entrepreneurial sector. We target a δ corresponding to the yearly rate of establishment exit recorded by the Business Dynamics Statistics survey from the U.S. Census Bureau, which averages 9.80% since the year 2000. We determine $\bar{\lambda}$ using the remaining endogenous part $\delta^{cd} = \bar{\lambda} R$ by setting the exogenous part $\bar{\delta}$ to 40% of the total rate Both parameters concerning access to entrepreneurship are set in such a way that mid-range ability workers have a chance to engage in research: $\psi = 1$ and $\bar{b} = 3$. Figure 1 plots both the share of households that have the opportunity and choose entrepreneurship given their type ν, i.e. $r(\nu)$, as well as the (scaled) distribution of ν in each occupation which results. For income volatility, we use Guvenen et al.'s (2021) estimate from a U.S. income dynamics study. We set the volatility of persistent innovations to $\sigma = 0.197$, the value estimated for their intermediate specification with Gaussian innovations and unemployment shocks that best resemble our model. Their study uses labor income data only but does include self-employment income attributed to labor. Theirs is a very large dataset, yielding a precise estimation and a good fit.

	Parameter	Value	Source
Population size	N	17.74	SCF
Discount rate	ρ	0.015	—
Labor supply elasticity	η	2	Chetty [2012]
Workers' ability dispersion	σ_ν	0.806	SCF
Exog. researcher entry rate	$\bar{\lambda}$	0.033	U.S. Census Bureau
Exog. entrepr. exit rate	$\bar{\delta}$	0.040	$\bar{\delta}/\delta = 40\%$
Barriers to entrepreneurship.			
Cost parameter	ψ	1	—
Opportunity parameter	$\bar{b}$	3	—
Productivity growth technology			
Curvature parameter	α	0	Akcigit et al. [2021]
Level parameter	β	0.039	SCF
Productivity/income volatility	σ	0.197	Guvenen et al. [2021]
Incumbent innovation parameter	ι	9.363	Garcia-Macia et al. [2019]
Researcher's effective labor	ξ	0.634	SCF
Final production parameter	θ	0.675	SCF
Intermediate production parameter	ϕ	0.482	$\phi = (1-\theta)/\theta$
Tax progressivity	τ	0.181	Heathcote et al. [2017]
Fraction of output taxed	ζ	0.200	Heathcote et al. [2017]
Growth step size	γ	1.083	Growth $= 2\%$

Table 1: Baseline parameter values

Income inequality. Using data from the Survey of Consumer Finances (SCF) we calculate five moments, averaged over seven surveys (those between 2001 through 2019), which are targeted for calibration. Those moments are: the Gini coefficient of income (0.56); the ratio of the 90th to the 10th percentile of income, P90/P10 (11.17); the share of income due to the top 1% of earners (19.8%); the share of income due to entrepreneurs (35.3%); and the share of entrepreneurs in the population (16.36%). To match the model's concept of entrepreneurship with the data we define it in the broadest sense: moments concerning entrepreneurship (the last two, for example) are computed for the model taking into account both researchers and established entrepreneurs. In the data, we consider entrepreneurs as those who reported being either self-employed or private business owners.[10] Using these targets we manage to replicate the data well along

[10]Despite working with the same data and similar categories, we adopt a different definition than that of Brüggemann [2021] or Cagetti and Nardi [2006]. We consider entrepreneurs as those households who

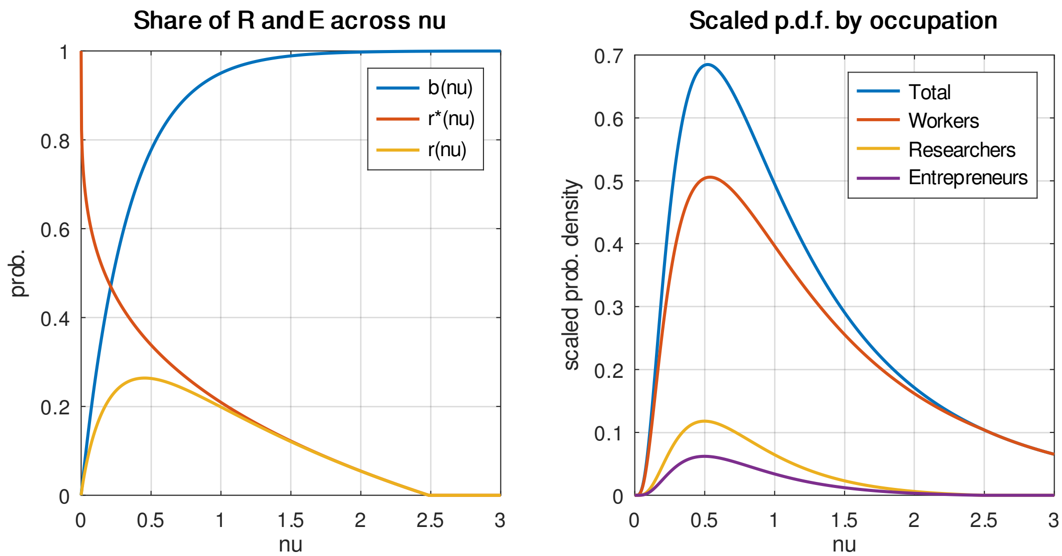

Figure 1: The entrepreneurial sector across ν: The left panel displays the share of the population for each type, ν in the three occupations: researcher and entrepreneur. The panel on the right displays the scaled (by the fraction of the population in each occupation) densities for each of the three occupations.

some relevant dimensions using the values for $N, \sigma_\nu, \beta, \xi$, and θ reported in table 1. Beyond the targeted moments, the model closely replicates some other income distribution features reported in table 2 and in the Lorenz curve plotted in figure 2. One feature the model does not succeed in replicating is the inequality among entrepreneurs, which it overshoots. Figure 3 plots the distribution of income for all occupations, making it clear that entrepreneurship is characterized by either low or high income depending on the state one occupies, giving rise to a high disparity in the cross-sectional distribution. In this case, the model slightly understates the inequality across workers to replicate the Gini coefficient of the whole economy.

Demography and preferences. The time discount rate is set at a standard value, $\rho = 0.015$. The disutility-of-effort parameter is set at $\eta = 2$, generally consistent with estimates of the Frisch elasticity [Chetty, 2012]. As already mentioned, population size N is calibrated so that entrepreneurs comprise the population share of 16.4% observed in the data.

answered accordingly to either one of the following two questions: "Do you work for someone else, are you self-employed, or something else?", and "Do you (and your family living here) own or share ownership in any privately-held businesses, including farms, professional practices, limited partnerships, private equity, or any other business investments that are not publicly traded?"

	Data	Model
Targeted moments		
Gini coefficient	0.566	0.566
P90/P10	11.17	11.17
Top 1% share of income	19.79	19.79
Entrpr.'s share of income	35.27	35.27
Entrpr.'s share of population	16.36	16.36
Non-targeted moments		
Bottom 5% share of income	0.39	0.47
Top 5% share of income	35.72	34.60
Top 10% share of income	46.15	45.74
Entrpr.'s Gini	0.646	0.767
Workers' Gini	0.494	0.430

Table 2: Moments of the income distribution: data vs. model

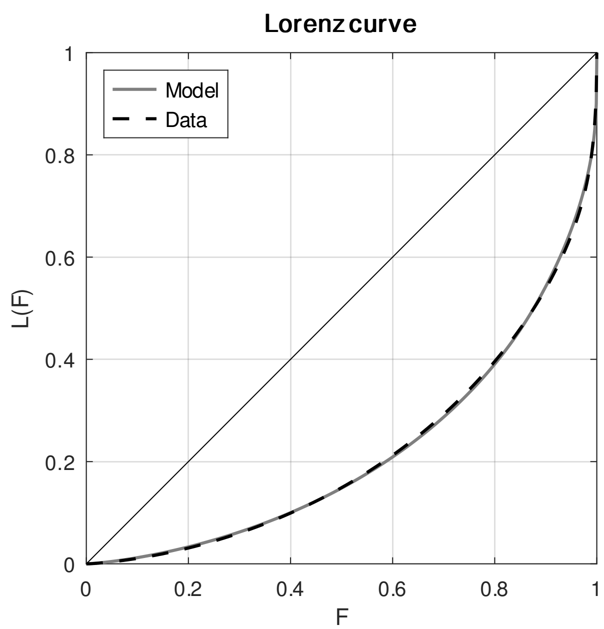

Figure 2: Lorenz curve for income: data vs. model

Taxation and growth. The baseline tax progressivity parameter $\tau = 0.181$ is taken from the Heathcote et al.'s (2017) estimation for the statutory tax rates in the U.S. using

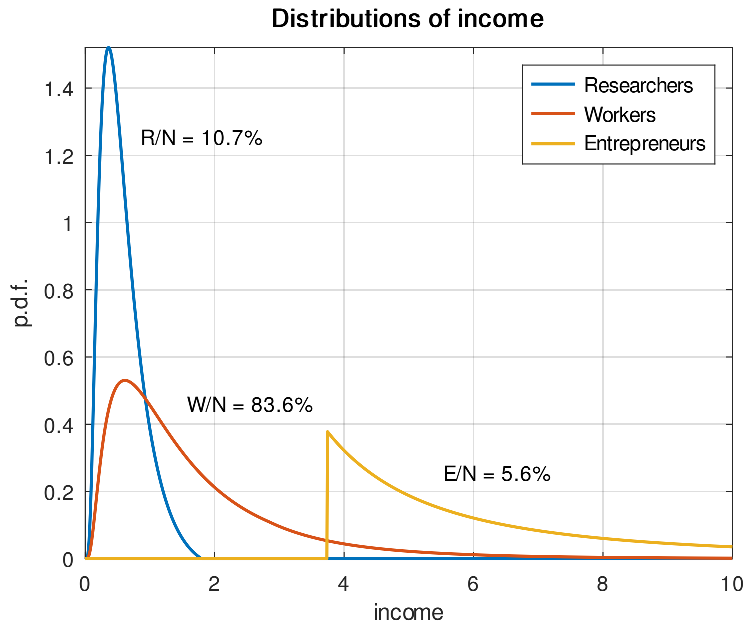

Figure 3: Distributions of income (before taxes) as well as the fraction of the population in each occupation.

the same functional form for the tax schedule we use here.[11] Similarly to these authors – see Jones [2022], too – we assume a fixed fraction $\zeta = 0.2$ of GDP is raised as tax revenue. Finally, concerning innovation and growth we need to set ι and γ. The former we calibrate so that the innovation due to creative destruction is one-fourth of the total [Garcia-Macia et al., 2019] and the latter is set so that the economy grows at a rate $g = 0.02$.

Elasticity of incumbent innovation. We assume in the baseline parametrization $\alpha = 0$ – a linear relation between entrepreneurial effort and productivity growth and innovation. This parameter plays an important quantitative role in how these variables respond to taxation, which in turn bears heavily on the results. Akcigit et al. [2021] performed a detailed empirical study of the relationship between innovation and taxation and found that the patent count of individual innovators can be quite elastic to marginal income keep rates. Their estimates go as high as 0.82, but we use a more conservative elasticity of 0.31, found in Table C.37 of the appendix. This estimate is based on marginal tax rate

[11]Their estimation fits very well the actual distribution of pre- and post-government income ($R^2 = 0.91$).

variations at the 90th income percentile. As a proxy for innovation, innovators' patent count can be represented by μ, so that with $\alpha = 0$ our model produces an equivalent elasticity of 0.25. As a robustness check, we show in section **??** how other values for α influence the results.

III.2 Baseline results

In this section, we characterize quantitatively the BGP equilibria of our economy under different policies. Economies are normalized to period $t = 0$ so they are comparable in terms of the existing technology. All variables are at their long-run level which characterizes the BGP equilibrium. We are, therefore, abstracting from questions about transitions between equilibria since the model would become intractable.

III.2.1 Occupational choice

The first thing to consider is how progressivity affects entrepreneurial participation. There are a few factors that play into this. Since participation is aggregated up from (ν, κ), which is also, of course, determined by the underlying distributions, we can begin to investigate occupational choice by looking at those different types of households.

Consider what happens when we go from $\tau^{US} = 0.18$ to $\tau = 0.5$. Figure 4 plots a comparison between both economies in terms of the proportion of ν-households who choose entrepreneurship, i.e., $r(\nu)$. In total, 2% of the population exits entrepreneurship, becoming workers. In the plot we see that this is mostly due to households with lower ν, in the range where the distribution is more concentrated, being discouraged from entrepreneurship. Above a certain threshold, entrepreneurship is encouraged by progressivity but the net effect is to reduce the entrepreneurial sector size.

To better understand this result, consider the net value of engaging in research that we get from combining equations (32) and (33):

$$\begin{aligned}\rho\mathcal{D}(\nu,\kappa) &\equiv \rho V_0^R(\nu) - \rho V_0^W(\nu) - \kappa \\ &= \rho\mathcal{D}(1,\kappa) - \frac{\lambda}{\rho+\lambda+\delta}(1-\tau)\log\nu.\end{aligned}$$

First, note that the attractiveness of entrepreneurship is decreasing in ν. Establishing

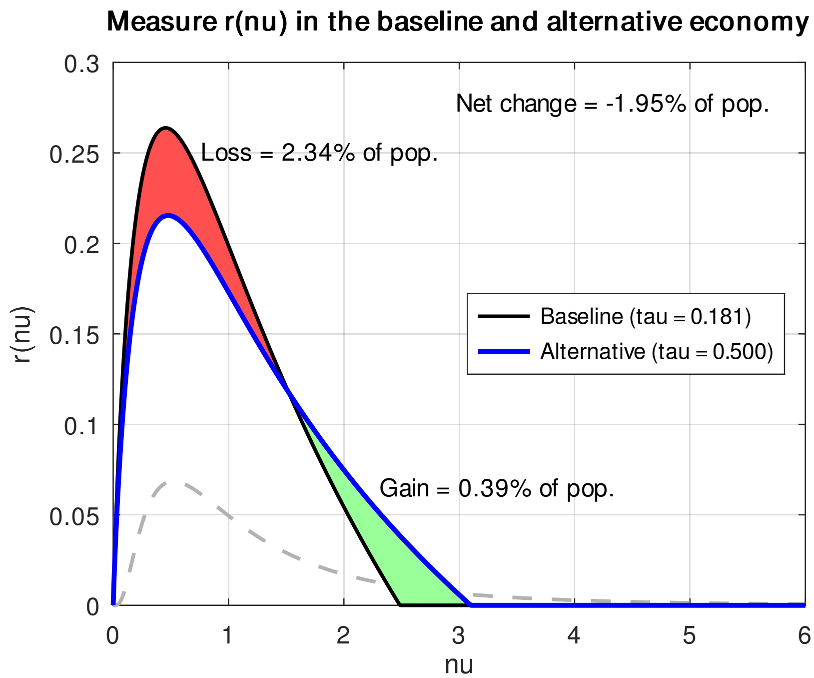

Figure 4: Change in the measure of households who pursue entrepreneurship given their type, $r(\nu)$: alternative vs. baseline economy. The dashed grey line traces the shape of the distribution of ν (not to scale).

oneself as an entrepreneur is the only way to become independent of one's work productivity. This makes entrepreneurship relatively more advantageous to low-productivity workers. The productivity impact is weighted by the time they will spend *not* working — hence, the term $-\lambda/(\rho+\lambda+\delta)$, being related to time spent in entrepreneurship — and by how much their productivity would affect their consumption as workers after taxes and transfers are applied. The more progressive and distributive the tax schedule, the less ν matters, which is expressed through $(1-\tau)$. Participation incentives are increasing in ν, but the relative impact is attenuated by progressivity.

The impact of an increase in τ on the net value of research can be decomposed into three effects: mechanical, behavioral, and general equilibrium effects. Mechanically, more redistribution and insurance reduces the disparity in consumption in all possible states. Then, there is a behavioral downward adjustment of labor supply and occupation that we described in the previous paragraph. Finally, there are new equilibrium prices and aggregates which further determine the comparative advantages of each occupation. Figure 9 in the appendix displays the relative importance of each effect at each productivity level, ν.

III.2.2 Characteristics of growth

Although participation in research monotonically decreases with rising progressivity, its contribution towards growth, i.e. creative destruction, as measured by the share $\delta^{cd}/\dot{n}$, displays a small variation – panel (a) of figure 5 – but reaching a maximum at around $\tau = 0.34$. Indeed, even if we consider a very wide range, $\tau \in [-0.2, 0.6]$, the variation in the share barely exceeds 1% and, even then, only when taxation becomes regressive. For more plausible values of τ, it varies within a 0.5% interval. This suggests that both types of innovation are more or less proportionally discouraged by progressivity.

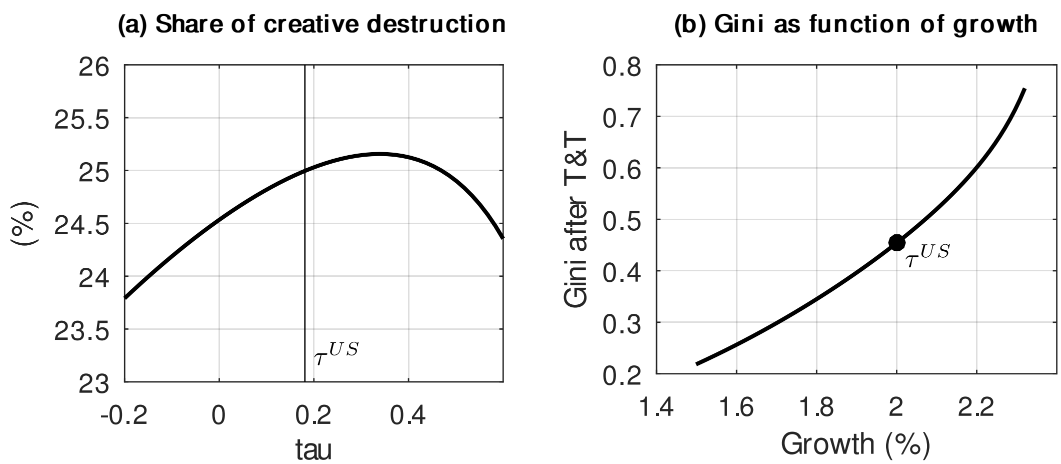

Figure 5: (a) Share of growth due to creative destruction: $\delta^{cd}/\dot{n}$. (b) Gini coefficient (after taxes and transfers) as a function of the growth rate; τ spans the interval $[-0.2, 0.6]$.

In panel (b), a quintessential trade-off of taxation between efficiency and equity is illustrated by the Gini coefficient of consumption as a function of the economy's growth rate. Naturally, through taxation, the planner can induce higher growth rates at the cost of more inequality by redistributing income less so that effort is more rewarded. One noteworthy feature, however, is the function's convexity; the higher the growth rate the more costly in terms of inequality additional growth becomes.

III.3 Optimal taxation

We start our normative analysis considering a utilitarian planner endowed with a social welfare function

$$\mathcal{U}_0 \equiv \mathbb{E}_h V_{h,0}. \tag{42}$$

To proceed, we introduce some additional notation. Let D_0 be the average present value of the disutility from effort and research sacrifice, κ, at time zero,

$$D_0 \equiv \mathbb{E}_h \left[\int_0^\infty e^{-\rho t} \left(\frac{\ell_h^{1+\eta}}{1+\eta} + \kappa \chi_h^E \right) \mathrm{d}t \right], \tag{43}$$

where χ_h^E is an indicator of h having chosen the entrepreneurship path. Let $\bar{c}_{h,t}$ be the certainty equivalent, along the constant growth path of consumption that delivers equilibrium value $V_{h,0}$ for household h, assuming equilibrium decisions ($\ell_{h,t}$ and choice of occupation) fixed. It is implicitly defined by

$$V_{h,0} = g/\rho + \log \bar{c}_{h,0} - \int_0^\infty e^{-\rho t} \left(\frac{\ell_{h,t}^{1+\eta}}{1+\eta} \right) \mathrm{d}t \tag{44}$$

and $\bar{c}_{h,t} = e^{gt} \bar{c}_{h,0}$; and let $\bar{y}_{h,t}$ be the corresponding pre-tax income necessary to attain it, i.e. $\bar{c}_{h,t} = \tilde{y}_t^{\tau} \bar{y}_{h,t}^{1-\tau}$. Finally, y_t with no h subscript denotes average income. With this notation, Proposition 5 offers a convenient way for writing $\mathcal{U}_0$.

Proposition 5. *In a BGP equilibrium, social welfare, $\mathcal{U}_0$, can be written*

$$\rho \mathcal{U}_0 = g/\rho - \rho D_0 + \log y_0 + \mathbb{E}_h \left[\log \left(\bar{y}_{h,0}^{1-\tau} \right) \right] - \log \left(\mathbb{E}_h \left[y_{h,0}^{1-\tau} \right] \right). \tag{45}$$

Equation (45) breaks down social welfare into easily interpretable terms. The first term represents welfare from economic growth; the second, disutility from effort and costs of research; the third, welfare from the time 0 output level of the economy; finally, the last two terms express the penalty the criterion applies to consumption disparity that remains after taxes and transfers. This latter component is comprised of permanent components due to heterogeneity and disparity across time in the form of income volatility that each household faces.

Figure 6 plots the variations from the baseline in social welfare, total and decomposed, as τ changes. The optimal progressivity is slightly lower than that of the U.S. tax system, with $\tau^* = 0.129$ as opposed to $\tau^{US} = 0.181$. Moving to the optimum results in a small improvement in welfare equivalent to a perpetual 0.46% increase in consumption to all households.[12] This suggests the current tax system is reasonable and easily justifiable either with slightly different parameters or with a planner who has a higher degree of inequality aversion (more on this later).

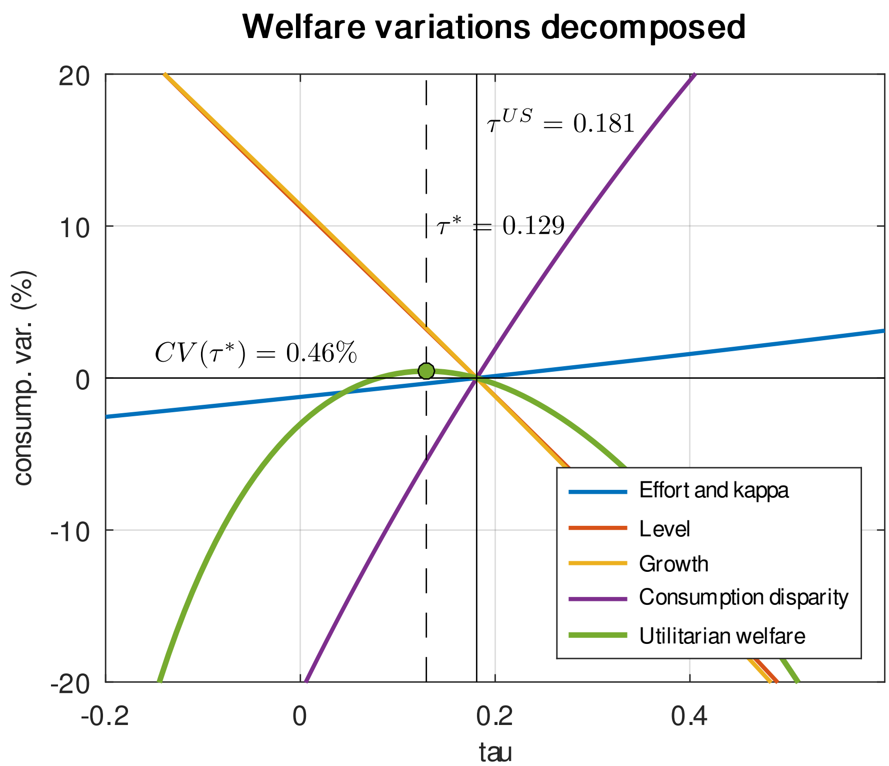

Figure 6: Welfare variations across τ decomposed into their components.

The contribution of each term in equation (45) to welfare changes is also displayed in Figure 6. Higher τ implies less overall effort and participation in research, resulting in less disutility from those sources. It also implies a very strong distributive positive effect on social welfare, representing the main reason for progressive taxation.

Arguably, however, the most crucial observation one can make is how surprisingly similar in magnitude the impacts on welfare coming from changes in both output level and growth are, with the red and yellow lines almost overlapping.

[12]Due to log-utility, the percentage in equivalent consumption variation can be calculated simply by finding the $CV(\tau)$ that equates $\log(1 + CV(\tau))$ to the difference in utility units between both equilibria. This can, of course, be done separately for every additive term in the social welfare equation.

III.4 The importance of growth

Consider an alternative economy that is identical to the baseline economy except for the fact that the growth rate is exogenous and kept fixed at $g = 0.02$. This eliminates any consideration for the impact of taxation on growth so that the plot in figure 6 would not have the yellow line. As we have just seen on that plot, disincentives to innovation can be as detrimental to welfare as are disincentives to production, so when those are removed we have a very different picture of the impact of taxation.

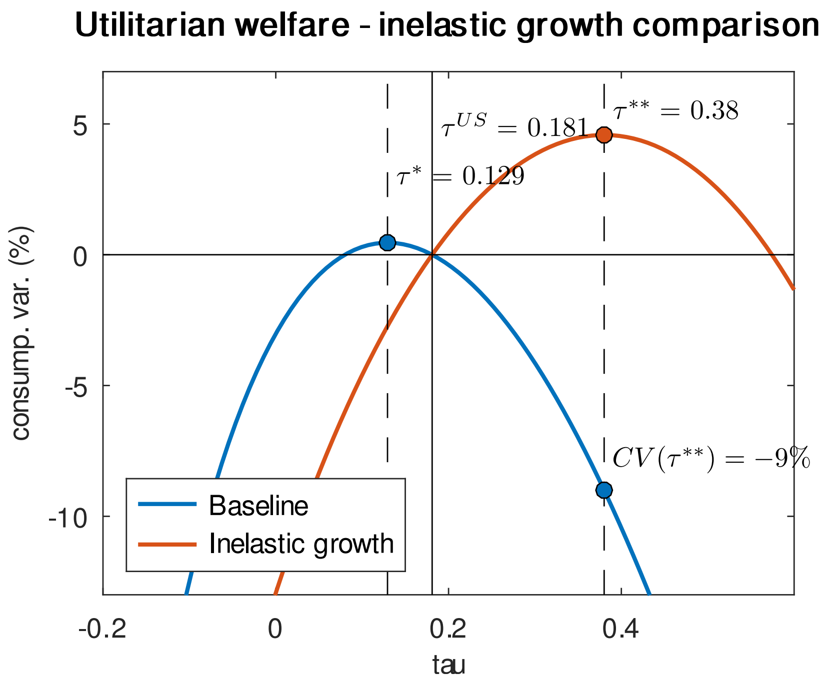

Figure 7: Welfare variations when comparing economies with elastic and inelastic growth.

Figure 7 shows how such an economy would compare with our baseline. In this inelastic growth case, the optimal progressivity is much higher, at $\tau^{**} = 0.38$. This translates into a marginal tax rate of 71% for the top income percentile, very close to policy recommendations that authors such as Diamond and Saez [2011] find using static models. This contrasts with a 47% rate implied by the current progressivity level and with a 39% by the optimal one with endogenous growth. Therefore, absent the growth mechanisms, our model manages to corroborate such results from the top income taxation literature.

The question is what happens when a planner acts on the assumption that impacts on growth are not relevant when in fact they are. Back in figure 7 we see that in our baseline scenario, a progressivity of $\tau^{**} = 0.38$ would result in an enormous loss of welfare equivalent to a decrease of 9% in perpetual consumption across households.

III.5 Robustness and Extensions

As with any study about taxation, quantitative findings are sensitive to assumptions about the values of elasticities. Whenever possible we simply adopt the conventional values from the literature. However, there is at least one crucial elasticity that we cannot simply borrow an agreed-upon value from the literature: that of mean productivity growth μ to policy τ. Needless to say by governing both long-run aggregate productivity, X, as well as the largest portion of innovation this elasticity is crucial for our analysis.

Recalling equations (22) and (23), we can see that

$$\frac{\mathrm{d}\log\mu(\ell^E)}{\mathrm{d}\log(1-\tau)} = \frac{1-\alpha}{\eta+\alpha}. \tag{46}$$

One way we can test different elasticities of output and innovation is by varying α, which we assume to be zero in the baseline. We can motivate a range of values for α using the research of Akcigit et al. [2021], as we have discussed in section III.1. Leaving aside innovation from creative destruction, which is responsible for about 25% of total innovation, we can use μ for calculating some measure of the elasticity of innovation to taxes. The authors find large elasticities, of up to 0.82, of patent counts to personal marginal rates faced by individual inventors. We use results from marginal keep rates at the 90th income percentile, reported in Table C.37 in their appendix, which produces more moderate elasticities. In a simple specification, the overall impact is a 0.31% increase in the patent count for every 1% increase in the keep rate. Controlling for inventor productivity changes the impact to 0.66% if the inventor is of high productivity, versus 0.18% otherwise.

This gives us a reasonable range of values to work with. Let y^{90} be the 90th percentile of income of the economy, and note that, after some algebra, we can write

$$\frac{\mathrm{d}\log\mu(\ell^E)}{\mathrm{d}\log(1-T'(y^{90}))} = \frac{\mathrm{d}\log\mu(\ell^E)}{\mathrm{d}\log(1-\tau)}\cdot\frac{1}{1+(1-\tau)\log(y^{90}/\tilde{y})}, \tag{47}$$

Table 3: Results for different α's.

Result	Curvature parameter, α			
	-0.2	0	0.2	0.4
Elasticity of μ to $1 - T'(y^{90})$	.33	*.25*	.18	.13
Optimal policy, τ^*	.06	*.13*	.19	.24
Top 1% marg. tax rate (%)	28.8	*39.3*	48.0	55.1
Welfare gain				
$CV(\tau^*)$ (%)	2.8	*0.5*	0.01	0.6
$CV(\tau^{**})$ (%)	-11.0	*-9.0*	-6.7	-4.2

Note — For every α, all parameters were recalibrated to match the targeted moments. The elasticity of μ is calculated in the baseline economy, with $\tau^{US} = 0.18$. Optimality is assessed according to the utilitarian criterion $\mathcal{U}_0$. τ^{**} denotes the optimal τ under the indicated parametrization when growth is inelastic; from 1st to 4th column: $\tau^{**} = .35, .38, .40, .42$.

into which we substitute equation (46). Varying α from -0.2 to 0.4 we get elasticities that go from 0.33 to 0.13, with $\alpha = 0$ producing a 0.25 elasticity. This suggests the values we consider are well within reason and are not far-fetched according to the empirical evidence; if anything they are slightly conservative.

Table 3 presents some results under four different values for α, including the corresponding elasticities. Unsurprisingly, as μ — and therefore output and innovation — become more inelastic, the optimal progressivity level becomes higher due to the planner being able to redistribute income at a lower cost to the economy. On the third row, we see what each progressivity level effectively means for the top 1% income percentile, with optimal marginal taxes ranging from 28.8% to 55.1%.

Across the board, changing to the optimal policy does not entail huge social benefits, as measured by $CV(\tau^*)$. In fact, with $\alpha = 0.2$, we have almost no benefit, with the U.S. progressivity being very close to optimal. Equally robust is the result of massive welfare decline under the naive optimal progressivity under inelastic growth, reported in the last line. In the most favorable case, there is still a sizable loss of 4.2% in perpetual consumption.

Table 4: Optimal policy τ^* for different α's according to different criteria.

	Curvature parameter, α			
Criterion	-0.2	0	0.2	0.4
Revenue maximization	.39	*.41*	.42	.43
Social welfare; $\omega =$				
0 (Inequality-neutral)	-.03	*.00*	.04	.08
1 (Utilitarian)	.06	*.13*	.19	.24
2	.20	*.27*	.32	.37

Note — For every α, all parameters were recalibrated to match the targeted moments. The baseline is $\alpha = 0$.

III.6 Different criteria

We have so far used the utilitarian criterion as a measure of social welfare. In this Section we clarify the relevant trade-offs, as with growth versus inequality, to allow one to define preferences over equilibria with any adopted objective. Borrowing from Benabou [2002] we define a general criterion for welfare that makes explicit how equity concerns can be taken into account on top of efficiency and insurance motives, allowing for consumption inequality penalization to vary from none to infinite.[13]

The idea is to first evaluate consumption uncertainty for each household using their preferences for risk before aggregating consumption across households. Thus, using the same definition of uncertainty equivalent $\bar{c}_{h,t}$ in equation (44), consumption is first aggregated as follows:

$$\overline{C}_t^{\omega} \equiv \left(\int_h \bar{c}_{h,t}^{1-\omega} \mathrm{d}h \right)^{\frac{1}{1-\omega}}, \tag{48}$$

where $\omega > 0$ indexes the degree of inequality aversion of the planner. This alternative criterion is defined as

$$\mathcal{E}_0^{\omega} \equiv \int_0^{\infty} e^{-\rho t} \log \overline{C}_t^{\omega} \, \mathrm{d}t - D_0. \tag{49}$$

The households' preferences for risk and time are held fixed while the planner is allowed to have a degree of flexibility to assess the distribution of consumption. Notice that ℓ and κ are taken into account through D_0 in the same way as in $\mathcal{U}_0$. The next proposition

[13] Introducing an inequality aversion parameter in the value functions is another possibility. Among other things, this would include inequality in the planner's objective.

provides an analogous decomposition to the one for $\mathcal{U}_0$.

Proposition 6. *The $\mathcal{E}_0^\omega$ criterion can be written as*

$$\rho\mathcal{E}_0^\omega = \frac{g}{\rho} - \rho D_0 + \log y_0 + \frac{1}{1-\omega}\log\left(\mathbb{E}_h\left[\bar{y}_{h,0}^{(1-\tau)(1-\omega)}\right]\right) - \log\left(\mathbb{E}_h\left[y_{h,0}^{1-\tau}\right]\right) \quad (50)$$

It is important to note that the insurance motive is still present when $\omega = 0$. Yet, the optimal level of progressivity is indistinguishable from 0 at our baseline calibration, and, if we consider the empirically plausible value of $\alpha = -0.2$ the system becomes slightly regressive when $\omega = 0$. In contrast, for $\omega = 2$, progressivity is substantially higher for our baseline calibration, but still substantially lower than what one would prescribe if the impact of policy on growth were ignored.

IV Conclusion

We have explored the equity-efficiency trade-off due to progressive taxation in a world of endogenous growth. In our baseline calibration, we find that while the optimal level of progressivity for the U.S. tax system is lower than the current one, transitioning to the optimal system yields moderate welfare gains. However, if the growth consequences of progressivity were disregarded, the ensuing prescription of a substantially more progressive system would lead to significant welfare losses.

For the model to remain tractable we have not taken into account the potential heterogeneity in research ability. As Jaimovich and Rebelo [2017] has shown this may cause (even linear) taxes to have a non-linear impact on growth. This type of consideration would likely reinforce our concerns with increased progressivity. All in all, our findings suggest caution in combating inequality with very progressive schedules.

References

Daron Acemoglu and Dan Cao. Innovation by entrants and incumbents. *Journal of Economic Theory*, 157:255–294, 2015.

Philippe Aghion and Peter Howitt. A Model of Growth Through Creative Destruction. *Econometrica*, 60(2):323–351, 1992.

Philippe Aghion, Ufuk Akcigit, and Peter Howitt. Lessons from schumpeterian growth theory. *American Economic Review*, 105(5):94–99, 2015.

Ufuk Akcigit, John Grigsby, Tom Nicholas, and Stefanie Stantcheva. Taxation and innovation in the twentieth century. *The Quarterly Journal of Economics*, 137(1): 329–385, jun 2021. doi: 10.1093/qje/qjab022.

Roland Benabou. Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, 70(2):481–517, mar 2002. doi: 10.1111/1468-0262.00293.

Bettina Brüggemann. Higher taxes at the top: The role of entrepreneurs. *American Economic Journal: Macroeconomics*, 13(3):1–36, July 2021. doi: 10.1257/mac.20170441.

Marco Cagetti and Mariacristina De Nardi. Entrepreneurship, frictions, and wealth. *Journal of Political Economy*, 114(5):835–870, oct 2006. doi: 10.1086/508032.

Raj Chetty. Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, 80(3):969–1018, 2012. doi: 10.3982/ecta9043.

Peter Diamond and Emmanuel Saez. The case for a progressive tax: From basic research to policy recommendations. *Journal of Economic Perspectives*, 25(4):165–190, 2011.

Xavier Gabaix, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. The Dynamics of Inequality. *Econometrica*, 84(6):2071–2111, 2016.

Daniel Garcia-Macia, Chang-Tai Hsieh, and Peter J. Klenow. How Destructive Is Innovation? *Econometrica*, 87(5):1507–1541, 2019.

Gene M Grossman and Elhanan Helpman. *Innovation and growth in the global economy*. MIT press, 1991.

Fatih Guvenen, Fatih Karahan, Serdar Ozkan, and Jae Song. What do data on millions of u.s. workers reveal about lifecycle earnings dynamics? *Econometrica*, 89(5):2303–2339, 2021. doi: 10.3982/ecta14603.

Jonathan Heathcote, Kjetil Storesletten, and Giovanni L. Violante. Optimal tax progressivity: An analytical framework. *Quarterly Journal of Economics*, 132(4):1693–1754, 2017.

Nir Jaimovich and Sergio Rebelo. Nonlinear effects of taxation on growth. *Journal of Political Economy*, 125(1):265–291, 2017.

Charles Jones. Taxing top incomes in a world of ideas. *Journal of Political Economy*, 130(9):2227 – 2274, 2022. URL https://EconPapers.repec.org/RePEc:ucp:jpolec:doi:10.1086/720394.

Charles I. Jones and Jihee Kim. A schumpeterian model of top income inequality. *Journal of Political Economy*, 126(5):1785–1826, 2018.

Tor Jakob Klette and Samuel Kortum. Innovating firms and aggregate innovation. *Journal of Political Economy*, 112(5):986–1018, 2004.

Wenli Li and Pierre Daniel Sarte. Progressive taxation and long-run growth. *American Economic Review*, 94(5):1705–1716, 2004. URL https://EconPapers.repec.org/RePEc:aea:aecrev:v:94:y:2004:i:5:p:1705-1716.

Sergio Rebelo. Long-run policy analysis and long-run growth. *Journal of Political Economy*, 99(3):500–521, 1991. URL https://EconPapers.repec.org/RePEc:ucp:jpolec:v:99:y:1991:i:3:p:500-521.

William J. Reed. The Pareto, Zipf and other power laws. *Economics Letters*, 74(1):15–19, 2001.

Artur Rodrigues. Optimal tax progressivity and growth, 2021. URL https://hdl.handle.net/10438/31408.

Paul M. Romer. Endogenous technological change. *Journal of Political Economy*, 98(5):S71–S102, 1990.

Joseph A. Schumpeter. *Capitalism, Socialism and Democracy*. Harper Perennial Modern Thought. HarperCollins, 3 edition, 1950.

A Derivations

Proof of Proposition 2. The final goods firm has to solve at a given time

$$\max_{\mathbf{q}} \left(\int_0^1 q_i^\theta \, \mathrm{d}i \right)^{\frac{1}{\theta}} - \int_0^1 p_i q_i \, \mathrm{d}i, \tag{51}$$

so that the first-order conditions for q_i yield the inverse demand curve

$$p_i(q_i) = \left(\frac{Y}{q_i} \right)^{1-\theta}. \tag{52}$$

The entrepreneurs solve, through their intermediate firms,

$$\max_{q_i} p_i(q_i)q_i - wL_i(q_i), \tag{53}$$

so that optimal pricing consists of the usual $1/\theta$ markup over marginal cost. When combined with (52), we can write

$$q_i = \left(\frac{1}{\theta}\frac{w}{\gamma^n x_i^\phi}\right)^{\frac{1}{\theta-1}} Y. \tag{54}$$

Now if we plug (54) into the final goods production function, using our assumption of $\phi = (1-\theta)/\theta$ and definition $X \equiv \int x_i \, \mathrm{d}i$, we get our final equation for wages:

$$w = \theta\gamma^n \left(\int_0^1 x_i^{\phi\frac{\theta}{1-\theta}} \, \mathrm{d}i\right)^{\frac{1-\theta}{\theta}} = \theta\gamma^n X^\phi. \tag{55}$$

To obtain the expression for Y, simply substitute (55) into (54), and then again into $L_i(q_i)$ and the aggregate labor definition $L \equiv \int L_i \, \mathrm{d}i$ to get

$$L = \frac{Y}{\gamma^n X^\phi} \tag{56}$$

which is the final form of Y.

At last, the expression for profits follows from substituting the expression for prices and wages into the profit equation,

$$\pi_i = \max_{q_i} \left(\frac{Y}{q_i}\right)^{1-\theta} q_i^\theta - \theta\gamma^n X^\phi \frac{q_i}{\gamma^n x_i^\phi}, \tag{57}$$

and then plugging the optimal q_i to get

$$\pi_i = (1-\theta)\gamma^n X^{\phi-1} L x_i \tag{58}$$

□

Proof of Proposition 3. Guess that the stationary distribution takes the form $f(x) =$

Cx^{-z-1} and plug it into (30) to get

$$
\begin{aligned}
0 &= -\delta^* C x^{-z-1} + z\mu^* C x^{-z-1} + z(z-1)\frac{\sigma^2}{2} C x^{-z-1} \\
&= C x^{-z-1}\left[-\delta^* + z\mu^* + z(z-1)\frac{\sigma^2}{2}\right].
\end{aligned}
$$

For the equation to hold for all x the term in brackets must be zero. Solving for the positive root we get z^* in (31) and $C = z^*$ follows from $f(x)$ integrating to one.

□

Constant growth. Observation: constant R_t is sufficient for constant growth. Remember that

$$\dot{n}_t = g_t / \log\gamma = \delta_t^{cd} + \iota\mu(\ell_t^E) = \bar{\lambda} R_t + \iota\beta(\ell_t^E)^{1-\alpha},$$

and therefore constant R_t, which is also sufficient for constant ℓ_t^E as stated in proposition 1, is all that is required for constant growth.

Proof of Proposition 3. Workers' value function. On a BPG, a ν-worker's value is given by

$$
\begin{aligned}
\rho V_t^W(\nu) &= gt + \frac{g}{\rho} + u_0^W(\nu) \\
&= gt + \frac{g}{\rho} + \tau \log \tilde{y}_0 + (1-\tau)\left(\log(w_0\nu) + \frac{\log(1-\tau)-1}{1+\eta}\right)
\end{aligned}
$$

where u_t^W denotes utility flow at t from optimal consumption and effort. To show this, simply note that $\tilde{y}_t$ and wages w_t grow at a constant rate, i.e. substitute $w_t = w_0 e^{gt}$, etc., substitute optimal effort $\ell^W = (1-\tau)^{\frac{1}{1+\eta}}$, and solve the integral in (1).

Entrepreneurs and researchers value function. From (16) we can establish the first-order condition for optimal entrepreneurial effort,

$$\ell_t^E(x_t, \nu) = \left(\beta(1-\alpha) x_t \frac{\partial V_t^E(x_t,\nu)}{\partial x}\right)^{\frac{1}{\eta+\alpha}}. \tag{59}$$

We guess $V_t^E(x,\nu) = V_t^{E_0}(\nu) + C\log x$, where, again, $V_t^{E_0}(\nu) \equiv V_t^E(x^0,\nu)$. Inserting the guess in equation (59) we get a constant optimal choice

$$\ell^E = (\beta(1-\alpha)C)^{\frac{1}{\eta+\alpha}}, \tag{60}$$

and the Ito calculus term defined in equation (17) reduces to

$$\frac{\mathbb{E}_t[dV_t^E(x_t,\nu)]}{dt} = C\tilde{\mu} + \frac{dV_t^{E_0}(\nu)}{dt}, \tag{61}$$

with $\tilde{\mu} \equiv \mu(\ell^E) - \sigma^2/2$. Then, inserting our guess into the left-hand side of (16) we can compare the coefficients of $\log x$ to get

$$C = \frac{1-\tau}{\rho+\delta}, \tag{62}$$

which is indeed constant in the BGP where R, and therefore δ, are constant.

Next, we form an ordinary differential equations system between V_t^R and $V_t^{E_0}$ in the BGP.

Suppressing the notation for type ν and denoting by u_t^R and $u_t^{E_0}$ their respective utility streams and using the fact that, if z is the equilibrium $f(x)$ shape parameter, $\log x$ is exponentially distributed with $\mathbb{E}[\log x] = 1/z$, the system can be written as

$$\begin{pmatrix} \dot{V}_t^R \\ \dot{V}_t^{E_0} \end{pmatrix} = \begin{bmatrix} \rho+\lambda & -\lambda \\ -\delta & \rho+\delta \end{bmatrix} \begin{pmatrix} V_t^R \\ V_t^{E_0} \end{pmatrix} - \begin{pmatrix} u_t^R + \frac{1-\tau}{\rho+\delta}\frac{\bar{\lambda}}{z} \\ u_t^{E_0} + \frac{1-\tau}{\rho+\delta}\tilde{\mu} \end{pmatrix}. \tag{63}$$

Note that, due to constant income growth, both utility streams are linear in time, specifically, $u_t = u_0 + gt$. So we use that functional form with undetermined coefficients to look for a particular solution for the system. Solving for the coefficients, we get

$$\rho V_t^R(\nu) = gt + \frac{g}{\rho} + \tau \log \tilde{y} + \frac{(\rho+\delta)v^R(\nu) + \lambda v^E}{\rho+\lambda+\delta} \tag{64}$$

$$\rho V_t^{E_0}(\nu) = gt + \frac{g}{\rho} + \tau \log \tilde{y} + \frac{(\rho+\lambda)v^E + \delta v^R(\nu)}{\rho+\lambda+\delta} \tag{65}$$

where $v^R(\nu), v^E$ are related to the value flows that are particular to the time spent in each occupation:

$$v^R(\nu) = (1-\tau)\left[\log(\xi w_0 \nu \ell^R) + \frac{\bar{\lambda}/z}{\rho+\delta}\right] - \frac{(\ell^R)^{1+\eta}}{1+\eta}, \tag{66}$$

$$v^E = (1-\tau)\left[\log(m_0) + \frac{\tilde{\mu}}{\rho+\delta}\right] - \frac{(\ell^E)^{1+\eta}}{1+\eta}. \tag{67}$$

□

Social welfare criteria. Let us consider two social welfare criteria: a utilitarian aggregation and an aggregate efficiency criterion due to Benabou [2002]. First, define D as the average household expected lifetime disutility from effort at time zero; $\bar{c}_{ht}$ and $\bar{y}_{ht}$ as, respectively, the certainty equivalent stream of consumption of household h and the corresponding pre-tax income necessary to attain it, i.e. $\bar{c}_{ht} = \tilde{y}_t^{\tau} \bar{y}_{ht}^{1-\tau}$, when holding expected disutility from effort unchanged. The criteria are defined as [14]

$$\mathcal{U} \equiv \int V_{h,0}\, \mathrm{d}h \tag{68}$$

$$\mathcal{E} \equiv \int_0^\infty e^{-\rho t} \log \bar{C}_t \, \mathrm{d}t - D, \tag{69}$$

where $\bar{C}_t \equiv \int \bar{c}_{ht}\, \mathrm{d}h$. Then, it can be shown that in a BGP the criterion is reduced to

$$\rho\mathcal{U} = g/\rho - \rho D + \log y_0 + \int \log\left(\frac{\bar{y}_{h0}}{y_0}\right)^{1-\tau} \mathrm{d}h - \log \int \left(\frac{y_{h0}}{y_0}\right)^{1-\tau} \mathrm{d}h \tag{70}$$

$$\rho\mathcal{E} = \underbrace{g/\rho - \rho D + \log y_0}_{\equiv \rho\mathcal{W}} + \underbrace{\log \int \left(\frac{\bar{y}_{h0}}{y_0}\right)^{1-\tau} \mathrm{d}h - \log \int \left(\frac{y_{h0}}{y_0}\right)^{1-\tau} \mathrm{d}h}_{\log\left(\frac{\int \bar{y}_{h0}^{1-\tau}\,\mathrm{d}h}{\int y_{h0}^{1-\tau}\,\mathrm{d}h}\right)} \tag{71}$$

where y with no household subscript is per capita income.[15] Due to Jensen's inequality, we can see that $\mathcal{U}$ is no greater than $\mathcal{E}$, and due to risk aversion, that $\mathcal{E}$ is no greater than $\mathcal{W}$. The presence of either risk or income inequality determines equality or inequality:

[14] Consider all integration over households normalized by population size: $\int \mathrm{d}h = 1$.

[15] When we present the welfare decomposition we use $\rho\mathcal{U} - \log y_0$ and $\rho\mathcal{E} - \log y_0$, instead of $\rho\mathcal{U}$ and $\rho\mathcal{E}$ to preclude the initial level of income, which is somewhat arbitrary, to affect the wrong attribution of

inequality	risk	
yes	yes	$\mathcal{U} < \mathcal{E} < \mathcal{W}$
no	yes	$\mathcal{U} = \mathcal{E} < \mathcal{W}$
yes	no	$\mathcal{U} < \mathcal{E} = \mathcal{W}$
no	no	$\mathcal{U} = \mathcal{E} = \mathcal{W}$

Certainty equivalent. Workers are not subject to risk, and therefore $c_t^W = \bar{c}_t^W$. As for researchers and entrepreneurs, we define

$$d^R = \frac{\rho + \delta}{\rho + \lambda + \delta} \quad \text{and} \quad d^E = \frac{\rho + \lambda}{\rho + \lambda + \delta},$$

and use the value functions to obtain the deterministic, constant growth consumption streams that attain the same utility,

$$\frac{\bar{c}_t^R(\nu)}{y_t} = \frac{\bar{c}_0^R(\nu)}{y_0} = \left(\frac{\tilde{y}_0}{y_0}\right)^{\tau} \left(\frac{\bar{y}_0^R(\nu)}{y_0}\right)^{1-\tau}, \tag{72}$$

where

$$\left(\frac{\bar{y}_0^R(\nu)}{y_0}\right)^{1-\tau} = \left[\left(\frac{N\theta\xi\nu}{\bar{\nu}^W + \xi\bar{\nu}^R} \exp\left\{\frac{\bar{\lambda}/z}{\rho+\delta}\right\}\right)^{d^R} \left(\frac{N(1-\theta)}{z/(z-1)} \exp\left\{\frac{\tilde{\mu}}{\rho+\delta}\right\}\right)^{1-d^R}\right]^{1-\tau},$$

and

$$\frac{\bar{c}_t^E(\nu)}{y_t} = \frac{\bar{c}_0^E(\nu)}{y_0} = \left(\frac{\tilde{y}_0}{y_0}\right)^{\tau} \left(\frac{\bar{y}_0^E(\nu)}{y_0}\right)^{1-\tau}, \tag{73}$$

where

$$\left(\frac{\bar{y}_0^E(\nu)}{y_0}\right)^{1-\tau} = \left[\left(\frac{N\theta\xi\nu}{\bar{\nu}^W + \xi\bar{\nu}^R} \exp\left\{\frac{\bar{\lambda}/z}{\rho+\delta}\right\}\right)^{1-d^E} \left(\frac{N(1-\theta)}{z/(z-1)} \exp\left\{\frac{\tilde{\mu}}{\rho+\delta}\right\}\right)^{d^E}\right]^{1-\tau}.$$

Here, we have also used the following facts:

$$w = \theta Y/L, \qquad L = \ell^W(\bar{\nu}^W + \xi\bar{\nu}^R), \tag{74}$$

$$m = (1-\theta)Y/X, \qquad X = z/(z-1). \tag{75}$$

different aspects of allocation changes.

B Additional Results

B.1 Research vs. work

Consider a change in policy to $\tau = 0.5$, forming an alternative economy. The movement along the extensive margin of occupational choice can either be favorable or unfavorable to research, being aggregated up from the choices of (κ, ν)-type agents with several factors coming into play. More specifically, those who are able and also choose to research are aggregated across κ to form the measure $r(\nu)$, which is then aggregated across ν to form R. These aggregations depend, of course, on the assumed distributions of κ and ν.

Figure 8 depicts what happens to $r(\nu)$ and the overall effect on research that the alternative policy would have compared to the baseline. More progressivity makes research less attractive to low ν agents and more attractive to agents above a certain threshold, making the proportion of researchers in the former group diminish, while growing in the latter. However, since the lognormal distribution of ν is more concentrated in those lower values, the loss outweighs the gain so that the net effect on total research is negative.

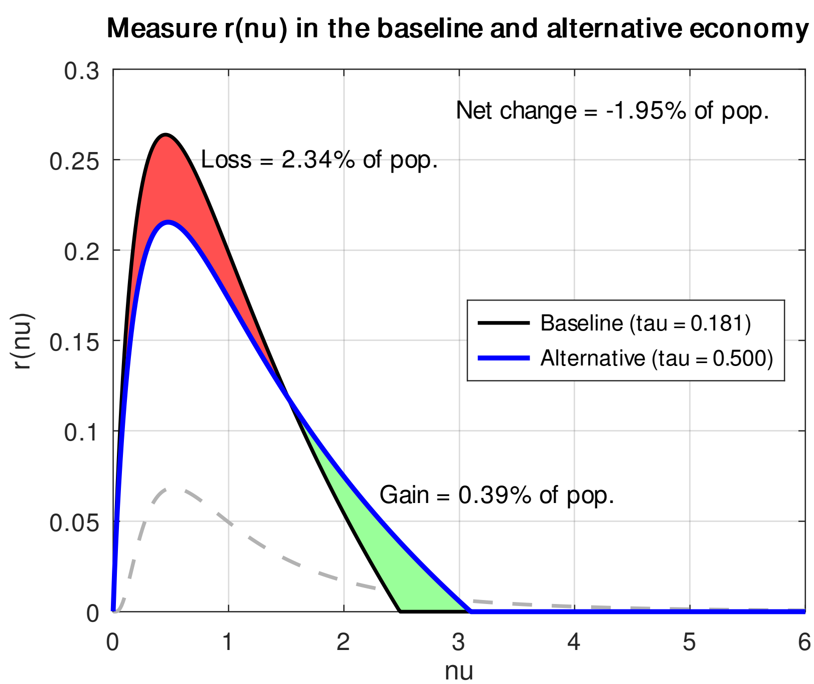

Figure 8: Change in the measure of researchers $r(\nu)$: alternative vs. baseline economy.

To gain some insight into the determinants of the comparative value of research, let us first define $\overline{\mathcal{D}}(\nu)$ to be the net value of entering research excluding the cost κ:

$$\begin{aligned}\rho\overline{\mathcal{D}}(\nu) &\equiv \rho\mathcal{D}(\kappa,\nu)+\kappa = \rho V_0^R(\nu) - \rho V_0^W(\nu)\\ &= \frac{\rho+\delta}{\rho+\lambda+\delta}(1-\tau)\left(\log(\xi)+\frac{\bar{\lambda}/z}{\rho+\delta}\right)\\ &+\frac{\lambda}{\rho+\lambda+\delta}\left[(1-\tau)\left(\log\left(\frac{m_0}{w_0\nu\ell^W}\right)+\frac{\tilde{\mu}(\ell^E)}{\rho+\delta}\right)-\frac{(\ell^E)^{1+\eta}-(\ell^W)^{1+\eta}}{1+\eta}\right].\end{aligned}$$

The first term on the right-hand side is the net value due to time spent in research, while the second pertains to time spent in entrepreneurship. Time in research differs from working in that effective labor supplied is lower — hence $\log(\xi)$ —, but there is also the expected value of entering entrepreneurship with some existing x — hence $\bar{\lambda}/z$. In entrepreneurship, there is the relative income advantage and expected productivity growth — hence the ratio $m_0/w_0\nu\ell^W$ and $\tilde{\mu}(\ell^E)$, respectively —, as well as the differences in effort. All of this is weighted taking into account the discounted probability of being in each state, hence the terms involving ρ, λ, and δ.

When the policy τ changes, there is a mechanical effect of smoothing consumption between both states through taxation, captured in the expression by $(1-\tau)$. Then, there is the behavioral effect of changes in optimal effort, ℓ^W, ℓ^E, to suit the new equilibrium. Finally, there is the general equilibrium effect of changes in prices m, w, in average entrepreneur productivity through z, and in entry and exit rates λ, δ through the new measure of researchers, R.

Figure 9 shows those effects stacked cumulatively and isolated for every type ν measured in equivalent changes in consumption. Examining the plot on the left we can see how the mechanical and general equilibrium effects interact with ν. Research for lower productivity workers is less attractive with higher τ since their low ν matters less and makes it not as desirable to be independent of it (through entrepreneurship) as when consumption is less smoothed. Since in equilibrium, the number of researchers is lower with more progressivity, the expected time in entrepreneurship is prolonged (through a lower exit rate), making the general equilibrium effect decrease with ν. Finally, due to log-utility, the behavioral effect does not interact with ν and is also smaller compared with the others for most agents.[16] When the effects are compounded, as shown in the plot on the right, the general equilibrium effect does not sufficiently compensate for the mechanical effect when varying ν, and so changes in research attractiveness remain increasing in ν, with low ν-agents being discouraged.

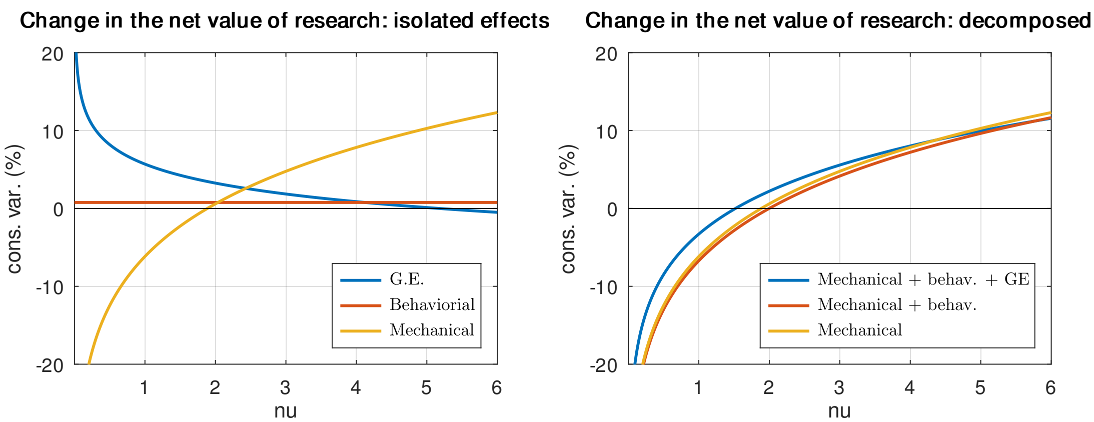

Figure 9: Decomposition of the changes in the net value of research: $\overline{\mathcal{D}}(\nu)$ in the alternative equilibrium ($\tau^{alt} = 0.5$) minus that of the baseline ($\tau^{US} = 0.181$). Values are in terms of permanent consumption variation. The left panel shows the changes of only one type of effect compared to the baseline; the right panel shows the cumulative changes of every effect type.

[16]By the envelope theorem, small changes in progressivity have almost no behavioral impact on utility, and therefore on occupational choice.