Shaping inequality: Progressive taxation under human capital accumulation

Danial Ali Akbari* Thomas Fischer†

February 6, 2024

This paper develops a model of human capital accumulation with on-the-job learning subject to obsolescence risk. The model analytically characterizes trade-offs of reforming the level of income tax progressivity and matches well income inequality in the US. An enriched version is used to quantitatively investigate optimal tax progressivity in the US. In contrast to standard models with exogenous income which suggest that progressivity is too low, accounting for endogenous human capital accumulation would call for lower tax progressivity. While optimal from a utilitarian perspective, such a reform mostly benefits individuals at the top 10% of the earnings distribution, however.

JEL classification: D31 – E24 – H21

Keywords: income inequality – fat tails – progressive income taxation – human capital accumulation

*Department of Economics, University of Oslo. Contact: d.a.akbari@econ.uio.no.

†Corresponding author. Department of Economics, Knut Wicksell Center for Financial Studies and Center for Economic Demography, Lund University Contact: thomas.fischer@nek.lu.se.

1. Introduction

Since the 1980s income inequality in the USA has been rising both at the top and the bottom end of the distribution (Piketty, 2014). While at the top end this can be traced back to the reduction of taxes (Piketty and Saez, 2007), the income at the bottom end is shaped by the risk of obsolescence. Due to various factors such as offshoring (Goos et al., 2014) or automatization (Hémous and Olsen, 2022) workers at the bottom end of the income distribution frequently experience their accumulated human capital becoming obsolete leading to a substantial loss in income. Under this backdrop, there is a call for the policy makers to reform the progressive tax-transfer system *shaping* this *inequality* (Heathcote et al., 2020b).

The design of an income tax-transfer system is always a trade-off between equity and efficiency. The reduced inequality due to an extended social insurance system has to be counter-weighed against the reduced labor supply which will result in efficiency losses. The literature (cf. e.g. the survey of Piketty and Saez, 2013) usually focuses on the immediate effect of reduced labor supply that reduces overall output and hence the aggregate size of the pie available for redistribution. However, there is a more long-run impact as progressive tax systems additionally discourage the accumulation of human capital which also shapes the income distribution.

We build a model of human capital accumulation *on the job* with uninsurable risks. In contrast to standard models in the Bewley-Huggett-Aiyagari tradition, income is modeled endogenously. More precisely, it is shaped by the individual labor supply decision, which subsequently shapes the accumulation of their human capital, and thereby the distribution of income by extension. This model not only captures the usual diffusion risk, but also jump risk associated with job destruction following labor-displacing automation or offshoring. Both of these events make certain human capital obsolete. Such obsolescence is crucial for realistic income distributions, especially at the bottom end. The calibrated model fits well the distribution of income in the US which is well approximated by Pareto tails both in the bottom and the top end (Toda, 2012).

This framework is then used to investigate the welfare impact of a progressive tax reform in the tradition of Heathcote et al. (2017). Despite the rich features, the model can be solved in closed form under the assumption of hand-to-mouth consumers. We show that the tax progressivity optimal under a utilitarian welfare function is substantially lower in a model with endogenous human capital accumulation as compared to a model with exogenous income. The concrete level increases with the sensitivity of labor supply to tax progressivity (Frisch elasticity) and the insurance motive of the social planner and is very sensitive to its choice.

In a second step, we enrich the presented framework using a model in the Bewley-Huggett-Aiyagari tradition in which individuals can react to adverse income shocks by both adjusting labor supply and forming precautionary savings. The model is solved in general equilibrium by means of a numerical algorithm which we develop and describe in Online Appendix F. In the extended model, the presence of (incomplete) financial markets provides more means for self-insurance and thus suggests lower optimal taxes as compared to the partial equilibrium model. In general equilibrium, more progressive taxes – as a measure of social insurance – also require less savings in physical assets and are associated with higher rates of return and lower wages. This is an implicit dividend for the asset-rich households going against the redistributive aim of the policy maker.

Our model calibrated to US evidence suggests an optimal regime with a progressivity of 9.6 %. This is around 50% lower than the status quo in the USA. This contrasts substantially to models that consider exogenous income suggesting higher optimal tax progressivity (cf. e.g. Kindermann and Krueger, 2022; Brüggemann, 2021). Our rich quantitative model, however, also reveals that while the flattening of the tax system encourages labor supply along the distribution, the net-gains in consumption of such a reform fall almost exclusively to the top 10%, making it questionable whether such a reform would find a democratic majority.

Literature: We substantially build on the framework of Heathcote et al. (2017) who introduce a parsimonious way of modeling a progressive income system and analytically characterize an optimal degree of tax progressivity in a steady state. In their general equilibrium model agents are subject to idiosyncratic wage shocks and can imperfectly insure themselves against the transitory (but not the permanent) shocks by means of labor supply or savings. Importantly, agents can once-for-all decide how much skills to accumulate (under ex-ante heterogeneity in learning ability) before entering the labor market which one can interpret as schooling akin to a Ben-Porath (1967) approach. In our model, we provide a novel perspective by considering *learning on the job* for which longer working hours not only immediately lead to higher earnings due to overtime payments, but also contribute to the accumulation of human capital and hence higher long-term earnings. This mechanism produces realistic endogenous earning patterns and emphasizes life-long learning in dynamic environments.[1] This has to be traded-off against a loss in leisure time.

In contrast, most of the literature (Huggett et al., 2011; Ludwig et al., 2012; Guvenen et al., 2014; Badel et al., 2020) using stochastic versions of Ben-Porath (1967) assumes that human capital is accumulated outside the job, affecting the budget constraint (learning has financial costs, e.g. college tuition) and even more importantly the time constraint (individuals that study reduce time spent working and hence their current income).[2] When analyzing fiscal policies with human capital, the existing literature thus emphasizes age-dependent taxes and incentives for formal schooling (Kapička, 2015; Stantcheva, 2017; Kapička and Neira, 2019; Heathcote et al., 2020b). Moreover, the literature usually does not focus on the macroeconomic and general equilibrium impact of said reforms. Optimal tax policies not only differ substantially between models with exogenous and endogenous human capital, but also crucially depend on the concrete form of human capital accumulation (on-the-job vs. outside-the-job) (Blandin and Peterman, 2019). In contrast to more general skills acquired with formal education outside of the job, our model emphasizes the firm- and job-specific human capital that is also subject to obsolescence risk once a work relationship is terminated. Jedwab et al. (2023) show that while the returns to job experience are lower than those to schooling, they are equally as important since individuals spend substantially more of their lifetime in market work than in formal schooling. Despite its empirical importance, on-the-job learning and its relationship with tax systems is only scarcely studied.[3]

Motivated by the empirical evidence (cf. e.g. Goos et al., 2014; Ebenstein et al., 2014; Autor et al., 2016), we also explicitly introduce the risk of human capital obsolescence[4]

[1]In order to jointly match the increase in earnings with age and the flat labor supply, Heathcote et al. (2020b) (building on Heathcote et al., 2017) need to assume that both exogenous efficiency levels and labor disutility are increasing in time. In our model the income and labor supply patterns in line with the evidence are endogenous model outcomes.

[2]While both features are present in the original work of Ben-Porath (1967), the more novel literature only features the latter.

[3]Peterman (2016) shows that accounting for learning-on-the job calls for substantially lower tax progressivity on labor income and instead higher taxes on capital income. His focus is, however, on life-cycle features and a simple piecemeal tax function with a single jump in marginal tax rates, while our model puts the emphasis on the cross-sectional distribution and using the more realistic Heathcote et al. (2017) tax function. In Peterman (2016), the optimal tax is associated with output losses due to reduced savings. We isolate the channel of changes in labor income tax progessivity and show that its reduction actually improves efficiency with the gains concentrated at the top end of the distribution.

[4]We think of *obsolescence* as exogenously arriving, while in Kredler (2014) individuals endogenously switch to a new technology in light of decreasing returns in old vintages.

shaping the bottom end of the income distribution. The underlying mechanism[5] gives rise to a double Pareto distribution in line with US evidence (Toda, 2012). Combining this with the complete income tax modeling in the tradition of Heathcote et al. (2017), we provide a comprehensive picture of both the income distribution and its taxation. In contrast, most of the literature focuses on taxes at the top end of the distribution (Diamond and Saez, 2011; Badel et al., 2020; Brüggemann, 2021; Kindermann and Krueger, 2022). Moreover, our investigation shows that while the flattening of the tax system encourages labor supply along the whole distribution, the net-gains in consumption of such a reform fall almost exclusively to the top 10%. This fact is disregarded in the literature focusing on revenue-maximizing tax reforms and lacking obsolescence shocks to capture the bottom end of the income distribution (Badel et al., 2020).

Finally, our paper also provides a contribution from the perspective of computational economics. The numerical solution of the model builds on earlier research (Nuño and Moll, 2018; Achdou et al., 2022) showing how to efficiently solve these models in a continuous-time setup. Given the highly non-linear nature of the considered problem here, we extend their algorithm by introducing the *pseudo-transient continuation* method in the considered economic setup.[6]

Structure: The remainder of the paper is organized as follows. Section 2 presents a tractable partial equilibrium model. In Section 3 we use the model calibrated to the US to conduct an analysis of optimal progressive taxes. An enhanced version of the model is presented in Section 4 also allowing for precautionary savings and considering general equilibrium effects. Section 5 considers optimal tax regimes in that setup. The final section provides some concluding remarks.

2. A model of human capital accumulation

We begin by discussing a partial equilibrium model in which individuals only have their labor supply as a margin of adjustment in light of adverse exogenous human capital shocks. This model is completely analytically tractable, allowing us to clearly identify the economic mechanisms, yet abstracts from some important facets. In Section 4, this is

[5]The mechanism was introduced in Reed (2001) and among others employed in the theoretical models of Aoki and Nirei (2017) and Jones and Kim (2018) that aim at explaining inequality in the USA.

[6]See Coffey et al. (2003) for a technical introduction to the pseudo-transient continuation method and Gomez (2015) for an application to economics.

supplemented by a general equilibrium model in the Bewley-Huggett-Aiyagari tradition in which individuals also have financial savings to buffer against adverse shocks.

2.1. Individual optimization

Setup: We assume that individuals adjust labor supply ℓ and consumption c under the following additive utility function:

$$U(c,l) = u(c) + v(\ell) = \ln(c) - \chi \frac{\ell^{1+\frac{1}{\upsilon}}}{1+\frac{1}{\upsilon}}, \tag{1}$$

for which $\upsilon > 0$ represents the Frisch elasticity of labor supply and $\chi > 0$ scales this effect as compared to the utility coming out of consumption. Due to the logarithmic preferences for consumption (unity in relative risk aversion), income and substitution effects here cancel out each other perfectly. This particular specification of the utility function was also employed in Heathcote et al. (2017). In the extended version of the model in Section 4, we explore the case with an intertemporal elasticity of substitution that suggests a dominating income effect.

Tax system: Pre-tax income $\tilde{y}_t = w \cdot h_t \cdot \ell_t$ depends on the amount of supplied labor ℓ, the human capital h, and the skill-based wage rate w. We assume a progressive tax system in the tradition of Bénabou (1996), with a tax progressivity τ_y. The latter models a progressive tax and transfer system in which marginal taxes are larger than average taxes. This modeling type has some convenient analytic properties and can be well fit to the US evidence (Heathcote et al., 2017). For the analytic characterization in this section, households are assumed not to form savings, the consumption thus equaling the post-tax income y:

$$c_t \equiv y_t = (w \cdot h_t \cdot \ell_t)^{1-\tau_y} T. \tag{2}$$

The scaling parameter T is chosen in order to guarantee a balanced budget of the government.

Human capital accumulation: Time is continuous and human capital h of heterogeneous individuals evolves according to the following process:

$$\begin{aligned} dh_t &= (A \cdot \ln(\ell_t) - \delta_h) h_t \cdot dt + \sigma \cdot h_t \cdot dZ_t + (h_0 - h_t) dJ_t \\ &= m_t \cdot h_t \cdot dt + \sigma \cdot h_t \cdot dZ_t + (h_0 - h_t) dJ_t. \end{aligned} \tag{3}$$

First, there is an idiosyncratic stochastic component with a standard deviation σ in the human income formation process. Second, households are exposed to sudden resets or jumps (dJ) with a Poisson arrival rate p. Once the jump arrives, the current specific human capital becomes worthless and individuals are reset to a common level h_0, which we normalize at $h_0 \equiv 1$.

A reset from above, when human capital is larger than the reset level $h > h_0$, can be thought of as human capital becoming obsolete during a sudden shock. Such shocks are prominent in the economy. First of all, individuals can lose their human capital due to disability or other severe persistent health shocks preventing them from continuing to work in their current occupation. Secondly, a lot of human capital is also specific to an industry (Neal, 1995; Sullivan, 2010), firm (Lazear, 2009) or certain task (Gathmann and Schönberg, 2010; Autor and Handel, 2013). Even if individuals stay within their occupation, a change of industry or employer will have made their industry or firm-specific human capital obsolete. Task specificity pertains to standardized procedures that are either offshored or automated. For individuals with $h < h_0$, being hit by the reset shock, eventually improves their human capital and monetizes their skills due to standardization of some production process in line with the conceptual framework from Acemoglu et al. (2012).

It is important to distinguish this sudden jump from the regular and expected depreciation of human capital modeled by the depreciation rate $\delta_h > 0$. A dynamic economy requires constant learning from the individuals in order to catch up with latest trends in technology or legislation even if they stay within their profession or firm.

In our model, individuals can actively accumulate human capital. Here, we model the case of *learning on the job*. The accumulation of human capital depends on the labor supply $\ell \geq 0$, i.e., people working a lot improve their ability in their job. Thus, there is a double rent from supplying labor ℓ. Not only does it increase your overall pay mechanically by working longer hours, but it also contributes to higher human capital h, resulting in a higher hourly skill-based wage. These advantages have to be counterbalanced against the general disutility of labor as captured by both the scaling parameter χ and the Frisch elasticity υ. We assume that accumulating human capital by working more entails decreasing returns to scale and we model this with a log-function. An intuitive interpretation of this production function is that it captures over-time hours supplied. For the case of $\ell = 1$ (and the absence of regular depreciation δ_h), we have no growth of human capital ($m = A \ln(\ell) = 0$). Individuals working more than the normalized level $\ell = 1$ thus furthermore accumulate human capital ($m > 0$), allowing them to increase their hourly payment. The scaling parameter $A > 0$ is considered the general efficiency of learning.

Optimization: Households discount time with a rate of time preference $\rho = \hat{\rho} + \psi > 0$ which can be decomposed in a pure time-preference rate $\hat{\rho} > 0$ capturing impatience and an age-invariant death probability ψ. In line with Blanchard (1985) we assume that a dying individual is replaced by a single child with human capital h_0 such that population size is constant. The objective of the individual is thus maximizing the present-value of its utility stream:

$$V \equiv \max_{c,\ell} \mathbb{E}_t \int_t^\infty \exp[-\rho(\tau - t)]U(c_\tau, \ell_\tau)d\tau, \tag{4}$$

subject to the the human capital accumulation process in (3), where V is the optimal current-value function. Consequently, we can derive the Hamilton-Jacobi-Bellman (HJB) equation (suppressing time indexes) characterizing the problem as follows:

$$\rho V = \max_{c,\ell} \left\{ U(c,\ell) + V_h(A\ln(\ell)h - \delta_h h) + V_{hh}\frac{1}{2}\sigma^2 h^2 + p(V(h_0) - V) \right\}. \tag{5}$$

The indexes indicate partial derivatives. The following proposition summarizes the individual optimal rules.

Proposition 2.1 (Optimal labor supply and consumption)
Optimal labor supply is given by:

$$\ell = \left(\frac{(1-\tau_y)(A+\rho+p)}{(\rho+p)\chi} \right)^{\frac{\upsilon}{1+\upsilon}} \geq 0. \tag{6}$$

Optimal consumption – which is equal to post tax income – amounts to:

$$y \equiv c = T(w \cdot h)^{1-\tau_y} \left(\frac{(1-\tau_y)(A+\rho+p)}{\chi(\rho+p)} \right)^{\frac{(1-\tau_y)\upsilon}{1+\upsilon}}. \tag{7}$$

In logs (allowing for a reasonable interpretation as elasticities) it is given by:

$$\ln(y) \sim \ln(T) + (1-\tau_y)[\ln(h) + \ln(w)] + \frac{(1-\tau_y)\upsilon}{1+\upsilon} \ln\left(\frac{(1-\tau_y)(A+\rho+p)}{\chi(\rho+p)} \right). \tag{8}$$

Proof of Proposition 2.1 The solution to the problem is suspended to Online Appendix A.

In general, the heterogeneous income is driven by the heterogeneous human capital (the second term in equation 8) whereas the other factors are common and just scale the income distribution relative to the human capital distribution. The previous equation will also be helpful to make a statement about welfare as it also reflects the utility out of consumption due to the assumption of a log-utility function in consumption. As the optimal decision ℓ is independent of the wage rate w, we normalize it to $w \equiv 1$ in the following without loss of generality.

The general process is displayed in equation (3) for which following from Proposition 2.1 the micro-founded drift-term m amounts to:

$$m = A\frac{\upsilon}{1+\upsilon}\ln\left(\frac{(1-\tau_y)(A+\rho+p)}{\chi(\rho+p)}\right) - \delta_h = m(\delta_h[-], p[-], \rho[-], \chi[-], \tau_y[-], A[+], \upsilon[+]). \tag{9}$$

Since there is a common labor supply ℓ, there is also a common drift m. The comparative statics presented in the brackets of the previous equation have some straightforward intuition. Not surprisingly, higher learning efficiency A increases labor supply and human capital accumulation. High depreciation of human capital – either in the form of regular depreciation δ_h or resets p – imply that human capital grows at a small rate. With respect to preferences, impatient individuals (high value of ρ) with strong disutility from labor supply χ supply less labor and thus experience a lower growth rate of human capital. Labor supply increases with the Frisch elasticity υ. For the extreme case of total inelastic labor supply ($\upsilon = 0$) the growth rate of human capital is independent of preference or policy parameters ($m = -\delta_h$). Most importantly from a policy perspective, a high tax progressivity τ_y discourages human capital investment.

2.2. Distribution of human capital, pre-tax and post-tax income

Distribution of human capital: The microfounded process of human capital is given in equation (3) with a constant drift as detailed in equation (9). This allows us to characterize the distribution of income in closed form in the following proposition.

Proposition 2.2 (The distribution of human capital)
The probability density function is given by:

$$f(h) = \begin{cases} \frac{\hat{C}}{h_0}\left(\frac{h}{h_0}\right)^{-(\vartheta_1+1)} & h \geq h_0 \\ \frac{\hat{C}}{h_0}\left(\frac{h}{h_0}\right)^{-(\vartheta_2+1)} & 0 < h < h_0 \end{cases}, \tag{10}$$

with an integration constant $\hat{C} = -\frac{\vartheta_1\vartheta_2}{\vartheta_1-\vartheta_2} > 0$ *and a mode* h_0*. The values* $\vartheta_1 > 0 > \vartheta_2$ *are given by:*

$$\vartheta_{1,2} = A_1 \pm \sqrt{A_1^2 + A_2}, \tag{11}$$

with:

$$A_2 = \frac{2p}{\sigma^2} > 0, \tag{12}$$

and:

$$A_1 = -\frac{m - \frac{1}{2}\sigma^2}{\sigma^2}. \tag{13}$$

A useful relationship between the values is $-\vartheta_1\vartheta_2 = A_2$ *and* $\vartheta_1 + \vartheta_2 = 2A_1$*.*

Proof of Proposition 2.2 The proof is given in Online Appendix B.

Recall that the process in (3) is a Geometric Brownian Motion that is reset to $h_0 = 1$ with an obsolescence (reset) rate of p. We require $\lambda_1 = p - m > 0$ for a finite mean. This is also the convergence rate to the mean which we employ as our measure of mobility. The resulting double Pareto distribution is characterized by both fat tails (with a tail coefficient $\vartheta_1 > 0$) at the right end (the *rich*) and at the left end of the distribution (the *poor*, tail-coefficient $\vartheta_2 < 0$). This structure was first discussed in Reed (2001). It has a high degree of tractability allowing to characterize several measures of inequality in closed-form including the top-shares and the Gini-coefficient as displayed in Table 1.[7]

It is easy to show that an increase in m decreases both the left tail $\vartheta_2 < 0$ as well as the right tail $\vartheta_1 > 0$. While the former implies a less fat tail in the lower end, the fat tails in the upper end (the top income earners) increases, implying a higher level of top inequality. Given that each individual supplies the same amount of labor (cf. Proposition 2.1), the distribution of human capital and the distribution of income coincide.

The distribution of post-tax income: The specific case of a progressive tax not only discourages human capital investment in the first place (lowering m), but furthermore puts a wedge between pre-tax income $\tilde{y}$ and post-tax income y. We can also exactly compute the distribution of income after taxes and transfers y as detailed in equation (2). In particular, we are able to compute a closed-form threshold of $T = y_{TF}^{\tau_y}$ making the tax system self-financing, where y_{TF} stands for the tax-free income. In other words,

[7]Note that both Heathcote et al. (2017) and Badel et al. (2020) impose a log-normal distribution with a fat Pareto-tail in the top end only, thus underestimating bottom inequality.

individuals earning a pre-tax income $\tilde{y}$ below this threshold ($\tilde{y} < y_{TF}$) are net transfer receivers, whereas individuals above this threshold ($\tilde{y} > y_{TF}$) are net tax payers.

In line with Heathcote et al. (2017) we assume that the surplus amount of taxes (Taxes $= \int_{y_{TF}}^{\infty}(\tilde{y}-y)df(h)$) relative to transfers (Transfers $= \int_0^{y_{TF}}(y-\tilde{y})df(h)$) constitutes government consumption G (Taxes $-$ Transfers $= gY = g\int_0^{\infty}\tilde{y}df(h)$) which represents a fraction $0 \leq g < 1$ of total output Y.[8] Following from a self-financing condition for a government without debt, this determines the threshold level of human capital for individuals that neither pay taxes nor receive transfers:

$$h_{TF} = h_0\left(\frac{(1-\tau_y-\vartheta_2)(1-\tau_y-\vartheta_1)}{(1-\vartheta_1)(1-\vartheta_2)}\right)^{\frac{1}{\tau_y}}(1-g)^{\frac{1}{\tau_y}}, \tag{14}$$

and correspondingly $y_{TF} \equiv \ell \cdot h_{TF}$. This level is strictly decreasing with the share g devoted to government consumption (e.g., the investment in public goods such as streets) and the degree of progressivity τ_y. For the special case of $g = 0$ all taxes would be transferred to other individuals with lower human capital in the form of transfers. For the special case of complete redistribution $\tau_y = 1$ we would have $h_{TF} = (1-g)H$, where $H \equiv \int_0^{\infty} hdf(h)$ is aggregate human capital.

It is easy to see that the relationship between the Pareto tail of pre-tax income $\tilde{y}$ and post-tax income y is given by:

$$\vartheta_{i,y} = \frac{\vartheta_{i,\tilde{y}}}{1-\tau_y} > \vartheta_{i,\tilde{y}}, \tag{15}$$

for $i = 1,2$ implying a more condensed distribution. The ratio of standard deviations of the log-measures is also given by $\frac{\sigma(\ln(\tilde{y}))}{\sigma(\ln(y))} = 1-\tau_y$.

3. Reforming the tax system

Calibration: In a next step, we take the micro-founded model to the data. We calibrate the model to US evidence for income from the 2013 Survey of Consumer Finances (SCF) reported in Kuhn and Ríos-Rull (2016).[9] Toda (2012) shows that the US income distribution can be well fitted by a double Pareto distribution as emerging from the process

[8] We use the convention of denoting aggregate measures by capital letters.

[9] We employ the measure of income (rather than earnings) provided in Kuhn and Ríos-Rull (2016). The latter also includes capital income and transfers. We include transfers to avoid bunching around nil income.

Measure	Var(ln(X))	Gini	Q_{90}/Q_{50}	Q_{50}/Q_{10}	Mean/Median	Share Top 1%
Data	1	0.58	3.33	3.46	1.85	19.7%
Model	1.04	0.58	2.9	3.5	1.8	23.0%

Table 1: Distributional measures: data (Kuhn and Ríos-Rull, 2016) vs. calibrated model ($\sigma^2 = 7\%$, $m = 2.8\%$, $p = 6.8\%$).
Closed-form expressions for all presented measures are presented in Online Appendix C.

detailed in this section.[10] We target values of both $\vartheta_1 = 1.5$ (Saez and Stantcheva, 2018) and $\vartheta_2 = -1.3$ to match the bottom end. Thirdly, we target a mean reversion pace $\lambda_1 = p - m = 4\%$ to match inter-generational correlation of income (Chetty et al., 2014).[11] The three moments (ϑ_1, ϑ_2, λ_1) are matched by the three structural parameters (p, m, σ) using the following relationships:

$$\begin{aligned}
\sigma^2 &= \frac{2\lambda_1}{2A_1 - 1 + A_2} > 0, \\
p &= \frac{A_2\lambda_1}{2A_1 - 1 + A_2} > 0, \\
m &= \frac{\lambda_1(1 - 2A_1)}{2A_1 - 1 + A_2}.
\end{aligned}$$

The fitted distribution matches both overall measures of inequality (Gini, Var(ln(X))), quantile ratios, and especially top shares (cf. Table 1). Note also that for the given distribution all the reported distributional measures can be expressed in closed form which adds tractability.

For the calibrated model the average annual growth is 2.8% – broadly in line with US evidence for growth of hourly wages in the life-cycle (Wallenius, 2011; Blandin, 2018) – with an annual variance of 7%. The expected growth rate is also broadly in line with the evidence in Jedwab et al. (2023) who document an average return to experience in developed economies of 2.6%-3.2% in their baseline results. The reset event emerges on average every $T = 1/p = 14.7$ years ($p = 6.8\%$). Without explicitly targeting, this result is broadly in line with the detailed micro evidence of Guvenen et al. (2021). Their econometric model features normal mixtures for both the permanent and

[10] Note that we do not directly employ the values reported in Toda (2012). He uses the residuals of the Mincer equation, implying that some of income distribution (especially at the top) is actually explained by the regression.

[11] With a generational length of 30 years we have a inter-generational correlation of $\exp(-\lambda_1 T) = 0.3$ in line with Chetty et al. (2014).

Variable	**Description**	**Value**	**Source**
δ_h	Depreciation human capital	1.5%	Guvenen et al. (2014)
ρ	Discount rate	5%	Krueger et al. (2016)
υ	Frisch elasticity	0.4	French (2005)
τ_y	Progressiveness income taxation	0.18	Heathcote et al. (2017)
g	Government share	0.189	Heathcote et al. (2017)
χ	Disutility labor	0.526	Calibrated to match $\ell = 1.4$
A	Learning efficiency	0.128	Calibrated to match $m = 2.8\%$

Table 2: Parametrization of the model.

The parameters are taken from the literature. The free parameter A is chosen in order to match the evidence for the drift term m reported in Table 1. The disutility of labor χ is set to match labor supply ℓ.

transitory shocks which the adverse case emerges with a probability of 5% and 11.8%, respectively.[12] Using the same process as suggested here, Jones and Kim (2018) also show how it matches well the evidence of leptokurtic income growth documented in Guvenen et al. (2021).

We also need to assign values to the parameters driving the economic decisions. For the parameterization we follow – if available – the existing literature to match the current state of inequality in the US given the prevailing tax rate τ_y. The values are summarized in Table 2.

We match a value of $\ell = 1.4$ which we interpret as a full-time job with the disutility of labor χ. Given the matched value of ℓ we can match m to the level suggested in Table 1 by setting the learning efficiency A. An individual that works 4 days a week (in model unit $\ell = 1.125$) would be characterized by a zero income growth (e.g., would not be promoted to senior positions).

Varying progressive taxation: Given the calibration aimed at fitting the US evidence we can make a counter-factual experiment regarding different (progressive) tax regimes by varying the parameter $\tau_y > 0$. As described in the comparative statics, higher taxation lowers the growth rate of human capital m (cf. left upper panel of Figure 1). Of course, increasing taxes reduces post-tax inequality. Before taxation the effect is slightly

[12] This case – closest to our considered case – is provided in specification 3 of Table IV in their published paper. In a working paper version they also featured the case with normal mixtures for the permanent shock only for which the adverse case emerged with a probability of 6.3% even closer to our case. They also explore more elaborate specifications featuring a large number of parameters that match the evidence even better. All of their specifications, however, lack a closed-form analytic expression strived for in our paper.

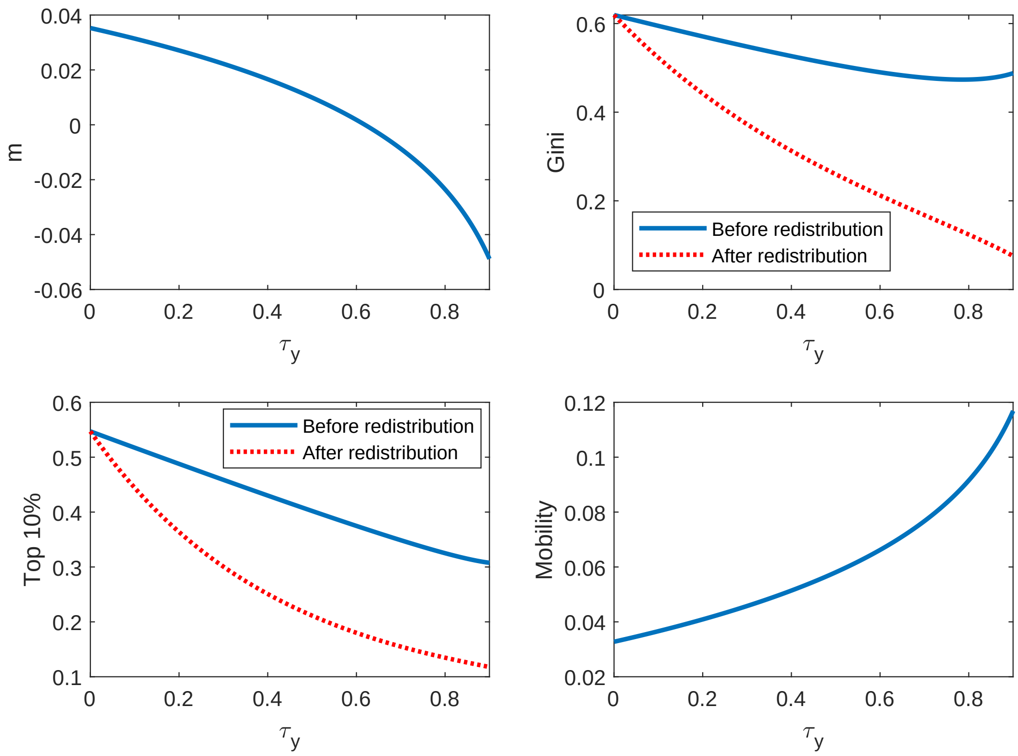


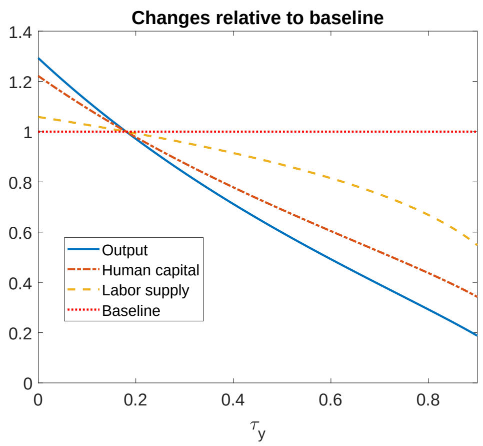
Figure 1: Comparative statics for a variation of taxation τ_y: Human capital growth m, Gini-coefficient, Top 10% income share, mobility.

Higher progressive taxation τ_y decreases the growth rate of human capital m and thus increases mobility. While the top-shares also decrease with taxation, the behavior is non-monotone for the overall Gini-coefficient. After redistribution inequality yet unambiguously decreases with taxation.

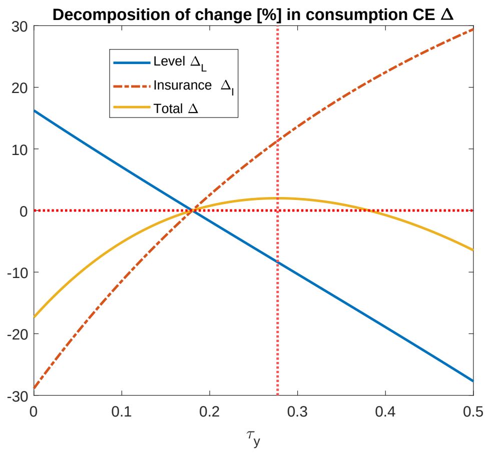
ambiguous. In fact, before redistribution we have a slight u-shaped response of inequality to an increase in taxation (see right upper panel of Figure 1). Due to higher taxes individuals supply less labor and also accumulate less human capital. For increasing taxation the left tail of the distribution becomes fatter as individuals accumulate less human capital, while the right tail – capturing top-inequality – is actually reduced. The share of income held by the top 10% (lower left panel of Figure 1)[13] strictly decreases with taxation. Of course, for maximum taxation ($\tau_y = 1$) we have total equality after redistribution.[14] Finally, taxation also impacts on mobility. The right lower panel in

[13]Here and in the following, we focus on the top 10%. Before taxes and transfers the model produces a share of 49%, even slightly overestimating the 47% reported in the data (Kuhn and Ríos-Rull, 2016).

[14]Note that we only consider a progressive taxation system ($\tau_y > 0$). Regressive systems can be modeled by $\tau_y < 0$. In fact, there exists a tax regime τ_y^* (which in our case is in the negative



(a) Aggregate measures



(b) Welfare

Figure 2: The effect of different tax regimes on aggregate measures (panel a) and the decomposed welfare effect (panel b).

Taxation τ_y lowers both immediate labor supply and human capital accumulation and thus average (total) income. For the given parameters the welfare-optimizing taxation is given by $\tau_y^{total} = 28\%$.

Figure 1 shows the mobility measure of mean reversion as Mob $= \lambda_1 = p - m$. Higher taxation lowers the growth rate of income m and thus increases mobility.

However, as usual the gain in equity comes at the cost of a loss in efficiency. As shown in Figure 2(a), an increase in taxes not only reduces labor supply but more so the accumulation of human capital and hence output. We normalize the value of average human capital to one for our benchmark calibration $\tau_y = 0.18$. Thus, a complete flat tax rate ($\tau_y = 0$) would increase overall (and average) output by approximately 30%.

Optimal taxes: In order to discuss optimal tax level we employ the concept of relative change in Certainty Equivalent (CE) consumption denoted by Δ. This is a common unit-free measure for the evaluation of policy changes (cf. e.g. Brüggemann, 2021). It accounts for the fact that both decision rules (c, ℓ) and equilibrium distributions ($f(h)$) change under different tax policies. More precisely, Δ measures the relative change to the baseline of CE consumption needed, for individuals to be at the same utility level after the shift in the tax regime from behind a proverbial veil of ignorance. Since this evaluation is carried out while individuals are ignorant about their standing, it is

domain), implying a growth rate of human capital m exceeding the destruction rate p ($m(\tau_y^*) > p$). In this case, we have a non-stationary distribution of exploding income inequality.

independent of their human capital, and thus a single constant value for every member of the population. Formally, Δ follows from:

$$\int \mathbb{E}_0 \left[\int_0^\infty \exp(-\rho t) U(c,\ell) dt \right] df(h) \Big|_{\tau_y'} = \int \mathbb{E}_0 \left[\int_0^\infty \exp(-\rho t) U([1+\Delta]c,\ell) dt \right] df(h) \Big|_{\tau_y^0}, \tag{16}$$

where τ_y^0 represents the baseline tax progressivity and τ_y' an alternative policy. For the concrete utility function assumed here it is given by:

$$\Delta(\tau_y') = \exp(W(\tau_y') - W(\tau_y^0)) - 1, \tag{17}$$

for which

$$W(\tau_y) = \int U(c,\ell) df(h) \Big|_{\tau_y} \tag{18}$$

represents the utilitarian welfare measure for some tax progressivity τ_y.[15] Figure 2(b) shows that maximum value for Δ emerges for a tax progressivity of $\tau_y = 28\%$ associated with a gain in CE consumption of roughly 2%. The level is substantially higher than the currently prevailing level of $\tau_y^0 = 18\%$. Even levels of tax progressivity as high as $\tau_y = 38\%$ are still associated with gains vis-à-vis the baseline scenario.

Welfare analysis: Similar to Heathcote et al. (2017) we can characterize the welfare function in closed form and thereby also have a deeper look at the countervailing effects of taxation on welfare. The complete welfare function (defined in equation 18) and its derivation is presented in Online Appendix D. This then allows us to analytically characterize the change in CE consumption Δ reported in Figure 2(b).

In the spirit of Flodén (2001) the change in CE consumption Δ can be decomposed into a (i) level effect Δ_L and (ii) an insurance effect Δ_I:

$$1 + \Delta = (1 + \Delta_L)(1 + \Delta_I). \tag{19}$$

[15] In the spirit of Heathcote et al. (2017) one might consider this as a *Ramsey*-style social planner problem. Heathcote and Tsujiyama (2021) argue that the outcome is close to the *Mirleesian* problem that is not constrained by a structural relationship in its tax policy instrument.

The level effect Δ_L evaluates changes in aggregate consumption C and changes in disutility related to labor supply and is formally given by:[16]

$$1+\Delta_L = \frac{C(\tau_y')}{C(\tau_y^0)} \cdot \exp\Big(v(\ell(\tau_y')) - v(\ell(\tau_y^0))\Big) = \underbrace{\frac{l(\tau_y')}{l(\tau_y^0)}}_{\text{Short run}} \cdot \underbrace{\frac{H(\tau_y')}{H(\tau_y^0)}}_{\text{Long run}} \cdot \underbrace{\exp\Big(v(\ell(\tau_y')) - v(\ell(\tau_y^0))\Big)}_{\text{Labor disutility}}. \tag{20}$$

The lower amount of hours worked ℓ under higher tax progressivity mechanically decreases the income available ($Y = C = HL = H\ell$) for consumption and thus utility out of consumption. This is associated with an immediate welfare reduction. Secondly, the lower supply reduces accumulation of human capital on the job and thus also reduces long-run human capital H. Once again this reduces income and consumption. Finally, higher progressivity reduces labor supply and hence reduces disutility associated with working thus improving welfare. However, the former two effects are dominating, implying that Δ_L is strictly decreasing with tax progressivity.

As displayed in Figure 2(a) this effect originates mostly from the long-run discouragement of labor supply and the associated loss in human capital. It is strictly decreasing with tax progressivity in a roughly linear manner (cf. Figure 2(b)). A representative agent considering the economy trough the lens of aggregate values and disregarding distributional issues would have as a goal to maximize Δ_L. For efficiency reasons, a flat tax $\tau_y = 0$ without transfers would be optimal.

With a heterogeneous population, however, this level effect Δ_L has to be counterbalanced against the insurance effect Δ_I, which in turn evaluates the changes in CE consumption ($\text{CE}[c(\tau_y)] = \exp(\int \ln(c) df(h))\big|_{\tau_y}$). It is formally given by:[17]

$$1+\Delta_I = \underbrace{\frac{\text{CE}[h^{1-\tau_y'}(\tau_y')]/H(\tau_y')}{\text{CE}[h^{1-\tau_y^0}(\tau_y^0)]/H(\tau_y^0)}}_{\text{Pure insurance}} \cdot \underbrace{\frac{h_{TF}^{\tau_y'}(\tau_y')}{h_{TF}^{\tau_y^0}(\tau_y^0)}}_{\text{Transfer}}, \tag{21}$$

which itself can be decomposed into a transfer effect and a pure insurance effect related to the certainty equivalent ratio of human capital after taxes.

High tax progression improves consumption insurance which increases utilitarian welfare. On the other hand, there is a *size-of-the-pie* effect with higher tax progressivity

[16]For a derivation refer to Online Appendix D.
[17]The detailed derivation is given in Online Appendix D.

	Endogenous HC	Exogenous HC ($A = 0$)
Utilitarian ($\tilde{\gamma} = 1$)	0.28	0.60
+Public Exp. ($\phi > 0$)	0.20	0.55
$\tilde{\gamma} = 0$	0.00	0.00
$\tilde{\gamma} = 2$	0.50	0.70
$\tilde{\gamma} \to \infty$	1.00	1.00

Table 3: Optimal tax progressivity τ_y for different welfare specifications.
Higher risk aversion ($\tilde{\gamma}$) of the social planner warrants higher tax progressivity. With the exception of a purely efficiency oriented social planner ($\tilde{\gamma} = 0$), optimal tax progressivity is larger than the observed status quo in the USA ($\tau_y = 0.18$). Preferences for government consumption ($\phi > 0$) eventually call for lower tax progressivity than the standard utilitarian approach. With exogenous human capital ($A = 0$), optimal tax progressivity under all preferences are substantially higher.

reducing overall output and hence the income available for transfers in the first place. The former effect, however, dominates (cf. Figure 2(b)) such that the the insurance measure Δ_I is strictly increasing with tax progressivity but at a diminishing rate. Given, however, that the level effect Δ_L falls almost linearly with tax progressivity, there exists therefore a local maximum for the total effect Δ that balances out these counteracting forces (cf. Figure 2(b)).[18]

Thus far, the welfare analysis was conducted under the (standard) premise of a utilitarian social planner. Outcomes are, however, sensitive to different assumptions about the social planner.

Preference for public expenditures: The share g used for non-transfer government expenses is not valued by the individual. One might consider this as the degree of the leaky bucket that is wasted in the public system when transferring resources from net tax payers to net transfer receivers. An increase in g reduces transfers and hence welfare. It is also important to note that this just acts as an offset and the optimal degree of tax progressivity τ_y is thus independent of the concrete level of g.

[18]Note that due to the hand-to-mouth assumption and the calibration to the income distribution, this exercise overstates consumption inequality and the insurance effect. We recalibrate the Pareto tails to the evidence for the consumption distribution reported in Toda and Walsh (2015) and assume them to be constant (for lack of a theory linking tax progressivity and the distribution of consumption). While the insurance effect is reduced, the adverse effect on aggregate income and consumption (similar to the case with $A = 0$) is ignored, eventually calling for a larger optimal $\tau_y = 0.36$. The extended framework in Section 4 allows for savings and hence provides a more quantitatively meaningful answer to the question of optimal taxation.

Let us contrast this to a scenario in which $G = g \cdot Y$ is consumption of public goods that is valued by the individual. We can introduce an additional homothetic preference for the public good consumption which weighted with a factor ϕ making individual utility:

$$U(c, \ell, g) = \ln(c) - \chi \frac{\nu \ell^{1+\frac{1}{\nu}}}{1+\nu} + \phi \cdot \ln(g \cdot Y). \tag{22}$$

As a result aggregate welfare with government consumption writes as $W_{gov}(\tau_y) \equiv W(\tau_y) + \phi(\ln(g) + \ln(\ell(\tau_y)) + \ln(H(\tau_y)))$. Following from the first-order condition ($\frac{\partial W_{gov}(\tau_y)}{\partial g} = 0$) it is easy to verify that the optimal level of government consumption is:

$$g = \frac{\phi}{1+\phi}, \tag{23}$$

reflecting the relative utility weight between individual and public consumption. Using the same reasoning, Heathcote et al. (2017) use the observed government share (18.9%) to back out the underlying preferences ($\phi = g/(1-g)$) which are positively related to each other.

The overall welfare function then reads:

$$W_{gov}(\tau_y) = W(\tau_y) + \underbrace{\phi(\ln(\ell(\tau_y)) + \ln(H(\tau_y)))}_{\text{Preference for public expenditure}} + \phi \ln\left(\frac{\phi}{1+\phi}\right). \tag{24}$$

In fact, a preference for the public good ϕ tends to promote lower taxes. The novel term depending on tax progressivity is strictly decreasing with τ_y since it reduces both immediate labor supply ℓ and long-run human capital H, i.e. it captures both of the negative effects associated with higher tax progressivity from the perspective of a representative agent. As such, a scenario with a preference for governmental consumption (i.e. $\phi > 0$) calls for lower tax progressivity. As shown in Table 3, in our calibrated model the optimal level of tax progressivity amounts to $\tau_y = 0.2$ lower than in the baseline and not very different than progressivity currently prevailing in the USA (i.e. $\tau_y = 0.18$).

The economic intuition is that the non-rival public good is available to all citizens alike. In particular, increases in such public expenditure will boost the utility level of the individuals at the bottom end of the distribution more substantially. For instance, even individuals with $h = 0$ (implying no income) will be able to consume the public good (e.g., can read all the books in the public library). As a result there is less of a need to provide individual consumption insurance by more progressive tax-and-transfer systems.

For large values of ϕ the welfare function is close to the one of the representative agent putting a lot of weight on aggregate output which – by assumption – will also increase the level of public consumption $G = g \cdot Y$.

Different welfare aggregators: The utilitarian aggregator assumes that the social planner and the individual have the same welfare function. Following the reasoning in Bénabou (2002) one can introduce a welfare function that allows the social planner to have a different relative risk aversion $\hat{\gamma}$ than the individual ($\gamma = 1$). Since in our framework labor supply is identical for all agents, so is the associated utility. Utility associated with consumption shall, however, be derived from the certainty equivalent level of consumption $\mathrm{CE}_{\hat{\gamma}}[c] = \left(\int c(h)^{1-\hat{\gamma}} df(h)\right)^{\frac{1}{1-\hat{\gamma}}}$ making the adjusted welfare function:

$$W_{\hat{\gamma}}(\tau_y) = \ln(\mathrm{CE}_{\hat{\gamma}}[c]) - \chi \frac{\upsilon \ell^{1+\frac{1}{\upsilon}}}{1+\upsilon}, \tag{25}$$

In Online Appendix D we show that in this case we have:

$$W_{\hat{\gamma}}(\tau_y) = W(\tau_y) - W_{ins,ut}(\tau_y) + W_{ins,\hat{\gamma}}(\tau_y), \tag{26}$$

for which $W_{ins,ut}(\tau_y) \equiv \ln(\mathrm{CE}[h^{1-\tau_y}])$ captures welfare effect of consumption insurance in the utilitarian case.

It can be verified that for $\lim_{\hat{\gamma}\to 1} W_{\hat{\gamma}}(\tau_y) = W(\tau_y)$ we get the utilitarian case. For $\hat{\gamma} \to 0$ we get the case of social planner who does not care about individual consumption insurance and only about aggregates $\mathrm{CE}_{\hat{\gamma}=0}[c] = C$:[19]

$$\lim_{\hat{\gamma}\to 0} W_{\hat{\gamma}}(\tau_y) = U(C(\tau_y), \ell(\tau_y)) + \ln(1-g) = U(H(\tau_y)\cdot \ell(\tau_y), \ell(\tau_y)) + \ln(1-g). \tag{27}$$

As such the optimization is the same as for the representative agent case generally suggesting low degrees of tax progressivity to minimize adverse effects on efficiency. In fact, in this situation the optimal tax rate is $\tau_y = 0$, i.e. a *laissez-faire* economy (cf. Figure 2(b)).[20] While there some positive effect coming from higher taxes that reduce labor supply and hence disutility of labor supply, the representative agent prefers an

[19] For a formal derivation refer to Online Appendix D.

[20] Note that we exclude the possibility of a regressive system $\tau_y < 0$ picking the corner solution as optimal. Regressive systems tend to look favorable for low $\hat{\gamma}$ and high ϕ. However, for the given calibration, negative τ_y tend to violate important technical conditions ($\vartheta_1 < 1 - \tau_y$) and are easily associated with non-ergodic distributions (explosion of inequality).

undistorted economy absent transfers. Here, the common average tax rate would be g in order to finance (non-transfer) government expenditures.

The diametrical extreme case is the one with $\hat{\gamma} \to \infty$ representing the Rawlsian social planner that only considers those at the bottom end of the distribution (i.e. those with $h \to 0$) and maximizes their consumption. The optimal tax progressivity here would be $\tau_y \to 1$.[21]

Of course, these mark the two extreme ends. In general, a social planner that has a stronger taste for consumption insurance as captured by larger values of $\hat{\gamma}$ will set higher rates of tax progressivity. The results are highly sensitive to the choice of this parameter. As reported in Table 3, already a value of $\hat{\gamma} = 2$ would for example imply an optimal tax progressivity as high as $\tau_y = 0.50$.

Exogenous human capital: A key innovation of this paper is to make the distribution of human capital – and thereby the distribution of income – endogenous. In our considered learning-by-doing setup, working more hours also contributes to a long-run increase of human capital which also boosts income and consumption possibilities.

If we, however, assume that there is no learning on the job ($A = 0$), the distribution of human capital and hence income is exogenously given and not affected by tax progressivity. Individual labor supply – still identical for all individuals and decreasing with tax progressivity – simplifies substantially to $\ell = \left(\frac{1-\tau_y}{\chi}\right)^{\frac{\upsilon}{1+\upsilon}}$. In this scenario both the Pareto tails ϑ_1 and ϑ_2 – characterizing the distribution – and the aggregate human capital H are independent of tax progressivity.

From the perspective of the representative agent considering levels Δ_L only (eq. 20) a *laissez-faire economy* with $\tau_y = 0$ would then be optimal. In contrast, only looking at the insurance part (Δ_I) would call for maximal progressivity $\tau_y = 1$ being optimal since there is no adverse long-run *size-of-the-pie* effect for higher tax progressivity.[22] Thus, even in this simplified scenario there is still a conflict between the insurance and efficiency motive of the social planner providing no unequivocal answer to the level of optimal taxation.

The model with exogenous human capital misses the long-run adverse impact on consumption through human capital and on the *size-of-the-pie* available for redistribution.[23]

[21] Note that it cannot be exactly $\tau_y = 1$ because this would imply no labor supply ($\ell = 0$) and hence also no output available for the social planner to redistribute.

[22] A formal derivation is presented in Online Appendix D.

[23] To ensure that the distribution is identical to the model with capital accumulation, we set a fixed drift term $m = 2.8\%$ as suggested in Table 1. This also reduces labor supply ℓ, which, following

As both effects call for relatively lower tax progressivity, a model with exogenous human capital would predict a higher optimal tax progressivity of $\tau_y = 61\%$. Thus, in line with Badel et al. (2020) and Wu (2021) – who, however, use learning-*or*-doing models à la Ben-Porath (1967) – the presence of human capital accumulation implies lower levels of optimal tax progressivity. In fact, a scenario with exogenous human capital accumulation is comparable to the problem under our endogenous human-capital-accumulation setting with a social planner that only cares about the immediate impact of its tax reform and disregards effects on long-run human capital accumulation covered in detail in Online Appendix E.

Ex-ante heterogeneity: The model thus far only features ex-post heterogeneity. Sometimes it is argued that earnings inequality is largely driven by ex-ante differences before entering the labor market (Huggett et al., 2011). It is straightforward to extend the analysis to account for discrete different learning abilities A_j.[24] For all practical purposes, we consider the most simple case of two different abilities associated with education: college (index c) with a population share of 47% in the adult US population, or high-school education and less (index h). Note that the approach is open to any discrete number of abilities. We calibrate $A_c = A \cdot (1 + \iota) > A = A_h = A/(1 + \iota)$ with $\iota = 18\%$ to match the substantial college premium of 82% (Heathcote et al., 2023); all other parameters unchanged. University graduates will work slightly longer hours and most importantly exhibit both larger income growth rates ($m_c > m_h$) as documented in the data (Cocco et al., 2005) as well as having higher average levels of human capital ($H_c > H_h$). Human capital for both types follows a double Pareto distribution ($f_j(h)$) with college graduates more likely at the top end of the distribution and vice versa for those with lower education. The progressive tax-and-transfer system will also redistribute between types with high school graduates being more on the receiving side. Formally, this will be captured in a pure redistributive measure Δ_R that is uniformly increasing in tax progessivity. Its quantitative impact is, however, neglectable and the usual insurance motive Δ_I remains dominant. Here, optimal tax progessivity amounts to $\tau_y = 25.1\%$, very close to the baseline result. Thus, the results are highly robust when allowing for ex-ante heterogeneity.

---

from the closed-form analysis, just shifts the welfare function downwards only changing its level but not its point of optimum τ_y^*.

[24] In line with Heathcote et al. (2017) one could also allow for preference heterogeneity χ_j with respect to labor supply disutility. Not only are those hard to observe, but they are of no concern for policy intervention.

Robustness checks: Diamond and Saez (2011) suggest an optimal marginal tax rate (MT) at the right top tail given by MT $= \frac{1}{1+\vartheta_1 \nu}$. Their approach also considers the top tail parameter ϑ_1 as exogenous and independent of the tax progessivity τ_y. For the given calibration ($\vartheta_1 = 1.5$ and $\nu = 0.4$), this calls for a value of 62.5%. Diamond and Saez (2011) report top marginal tax rates in the US (at the publication date of their paper) as 42.5% and thus at a substantially lower level.[25] For the top 1% in the baseline calibration our paper produces a marginal tax rate of 47% closed to the observed value and suggests an optimal value of 61%. Thus, the policy recommendation is broadly in line with Diamond and Saez (2011). Their approach is, however, computed under the assumption of tax revenue maximization for the government, thereby only taking into account the welfare of the individuals with the lowest income. In contrast, under identical preferences for all individuals their approach would suggest a zero top marginal tax rate. In comparison, our approach not only captures the top, but characterizes the complete system and is also less sensitive to changes in the social planner's objective function.

As usual, the results depend heavily on the parametrization of the Frisch elasticity υ. In Online Appendix D, we present exercises for either a lower Frisch elasticity ($\upsilon = 0.2$) or the case of unit elasticity (higher than the benchmark). In these exercises, in order to keep an labor supply ℓ unchanged relative to the baseline, an increase of the Frisch elasticity ν also requires an increase in the labor disutility χ. In general, optimal taxes decrease with the Frisch elasticity and are sensitive to assumptions on its value. The extreme case of inelastic labor supply ($\upsilon = 0$) would result in a complete redistributive tax ($\tau_y = 1$) being optimal even for a utilitarian case since there is no adverse reaction of the labor supply. A formal discussion is also found in Online Appendix D.

Thus far, we only considered long-run steady state outcomes. In Online Appendix E we also evaluate the transitional dynamics. While labor supply jumps down after changes in a rise in tax progressivity, human capital is only slowly de-accumulated. If the policy maker discounts for the transition, she has an incentive to set relatively higher levels of tax progressivity confirming earlier findings in the literature (cf. e.g. Bakis et al., 2015; Brüggemann, 2021). While the welfare gains from reducing inequality are immediately

[25] Formally, our approach even suggests $\lim_{y\to\infty} \text{MT} = \lim_{y\to\infty} 1 - (1-\tau_y)Ty^{-\tau_y} = 1$ and thus the highest possible level at the very top. This is regardless of the free parameter τ_y and follows from the specific parametric assumption in the tradition of Bénabou (1996).

reaped, its adverse consequences on aggregate human capital only manifest themselves in the long run making the social planner front-load welfare.

All in all, the model broadly suggests that optimal tax progressivity is substantially beyond what is currently in place in the US. This, however, might be the result of the strong assumption that individuals do not have access to savings to self-insure against labor income risk. In the following section, we relax this and other simplifying assumptions, and additionally consider a general equilibrium setup with private savings.

4. A full general equilibrium approach

To allow for analytic tractability the model presented so far has made some strong assumptions which are suspended in this section. We fuse the model presented so far with a rather standard Aiyagari (1994) model, in which individuals form precautionary savings under borrowing constraints to self-insure against idiosyncratic income risk with both labor and capital market clearing in general equilibrium. In contrast to Aiyagari (1994) labor supply is a decision variable and also endogenously determines human capital with learning-by-doing (LBD) which takes place on the job. We also allow for a realistic and rich tax system also featuring (linear) taxes on capital income and consumption. From now on, will refer to the previous setting explored in Sections 2 and 3 as the scenario with hand-to-mouth individuals or simply as the partial equilibrium (PE) model for brevity.

4.1. Extended setup

Private households solve the following HJB equation:

$$\rho V = \max_{c,\ell} U(c,\ell) + \frac{\partial V}{\partial h}\left(Aq(\ell) - \delta_h\right)h + \frac{\partial V}{\partial a}\underbrace{\Big((wh\ell)^{1-\tau_y}T + (1-\tau_K)ra - (1+\tau_C)c\Big)}_{s(a,h)}$$
$$+ \frac{1}{2}\sigma^2 h^2 \frac{\partial^2 V}{\partial h^2} + p(V(a,h_0) - V(a,h)), \tag{28}$$

for which $q(\ell)$ (with $q' > 0$ and $q'' < 0$) describes the accumulation of human capital on the job. Both capital income and consumption is also subject to taxation (with tax rates $\tau_K > 0$ and $\tau_C > 0$, respectively). Besides the usual first-order condition for consumption ($U_c = (1+\tau_C)\frac{\partial V}{\partial a}$), this also gives rise to a first-order condition for labor supply:

$$\underbrace{\frac{\partial V}{\partial a}(wh)^{1-\tau_y}(1-\tau_y)\ell^{-\tau_y}T}_{\text{I}} + \underbrace{\frac{\partial V}{\partial h}Ahq'(\ell)}_{\text{II}} = -U_\ell > 0. \tag{29}$$

This equation provides some key insights. It equates the marginal disutility from labor supply not only with the immediate rise the earnings (I), but also the long-run gains in human capital from working longer hours (II) providing a double dividend from working longer. The marginal gains of working longer for human capital are, however, decreasing in ℓ (due to $q'' < 0$). The same holds true for the immediate gains in a progressive tax system ($\tau_y > 0$) making (29) a highly non-linear equation that does not generally posses a closed-form solution.

Individuals cannot short assets ($a \geq 0$). They are, moreover, constrained by the total number of hours in a week ($0 \leq \ell \leq \ell_{max}$) for which ℓ_{max} is set to three full-time jobs. For the aggregate labor supply L, we target a full-time job ($L = \ell_{max}/3$). Following from the decision rules c and ℓ determining the accumulation of both physical assets a and human capital h, the population will be distributed on $f(a,h)$.

The representative firm uses aggregate capital K and human-capital augmented labor supply HL in a standard Cobb-Douglas technology maximizing profits:

$$\begin{aligned} &\max_{K,HL} Y - wHL - (r+\delta_k)K \\ &= \max_{K,HL} \Xi K^\alpha (HL)^{1-\alpha} - wHL - (r+\delta_k)K, \end{aligned} \tag{30}$$

where $0 < \alpha < 1$ is the capital share, $\Xi > 0$ is the total factor productivity (TFP) and $\delta_k > 0$ is the depreciation rate. The stationary equilibrium is then defined as follows.

Stationary equilibrium

1. For given efficiency-unit wages w and interest rates r, households of total unit measure choose consumption c and labor supply ℓ in order to maximize The HJB equation (28).

2. For given efficiency-unit wages w and interest rates r, the representative firm chooses capital K and labor-augmented human capital HL to maximize their profits according to equation (30) implying $r + \delta_k = \alpha\frac{Y}{K}$ and $w = (1-\alpha)\frac{Y}{HL}$.

3. For a given tax progressivity τ_y, consumption τ_C and capital income taxes τ_K as well as government consumption share g, the government sets T in order to guarantee a balanced budget:

$$G = g \cdot Y = \tau_K r K + \tau_C C + T \int_0^\infty \int_0^\infty (w \cdot h \cdot \ell(a,h))^{1-\tau_y} df(a,h) = \text{Taxes} - \text{Transfers}, \tag{31}$$

where $f(a,h)$ is the stationary joint distribution of physical assets a and human capital h given by the Kolmogorov-Forward equation

$$-\frac{\partial}{\partial a}\left[s(a,h)f(a,h)\right] - \frac{\partial}{\partial h}\left[\left(Aq(\ell) - \delta_h\right) h f(a,h)\right] + \frac{1}{2}\frac{\partial^2}{\partial h^2}\left(\sigma^2 h^2 f(a,h)\right) + p(f(a,h_0) - f(a,h)) = 0 \tag{32}$$

such that $\int_0^\infty \int_0^\infty f(a,h) dh da = 1$ and with the savings function $s(a,h)$ as in (28).

4. Capital markets clear – i.e. $K = \int_0^\infty \int_0^\infty a df(a,h)$ – and labor markets clear – i.e. $HL = \int_0^\infty \int_0^\infty h\ell(a,h) df(a,h)$. Moreover, according to Walras' law goods markets also clear:

$$Y = \int_0^\infty \int_0^\infty [c(a,h) + s(a,h)] df(a,h) + G + \delta_k K. \tag{33}$$

To numerically solve the model, we develop an algorithm using recent advances in continuous time methods (Nuño and Moll, 2018; Achdou et al., 2022) that is detailed in Online Appendix F.

4.2. Calibration and baseline results

The parametrization follows a common two-step procedure with some values being set in line with the literature and others calibrated to match chosen moments.

Calibration: For the calibration we broadly stick to the values already employed in the partial equilibrium exercise and detailed in Table 2. The exogenously set parameters differing from the ones in the partial equilibrium economy or not available are summarized in Table 4.

Not being constrained by the need of analytic tractability we allow utility out of consumption to have a dominating income effect – i.e. $u(c) = \frac{c^{1-\gamma}-1}{1-\gamma}$ with $\gamma > 1$ – in line with the macroeconomic literature (Brüggemann, 2021; Kindermann and Krueger, 2022). Furthermore, for the human capital production, we consider a more general

Category	**Name**	**Value**
Preferences	Risk aversion γ	1.1
Production	Capital share α	0.36
Taxes	Capital income tax τ_K	28.3%
	Consumption tax τ_C	5%
Human capital production	Curvature θ	0.75

Table 4: Exogenously set parameters additional to or different from the partial equilibrium model (summarized in Table 2).

The choice of parameters deviating from the ones in the partial equilibrium exercise is explained in the text.

Category	**Name**	**Value**	**Goal**	**Data**	**Model**
Preferences	Disutility labor χ	0.16	Average labor supply	1.4	1.40
Income process	Volatility σ^2	4.6%	Gini income	0.58	0.58
Production	Total Factor Productivity Ξ	0.29	Wealth / output ratio	3	2.99
	Depreciation δ_k	7.5%	Interest rate	4.5%	4.56%

Table 5: Calibrated parameters with targeted moment and fit.

The calibrated model fits the targeted moments very well. Average / aggregate labor supply equals on full-time position $L = \ell_{max}/3$.

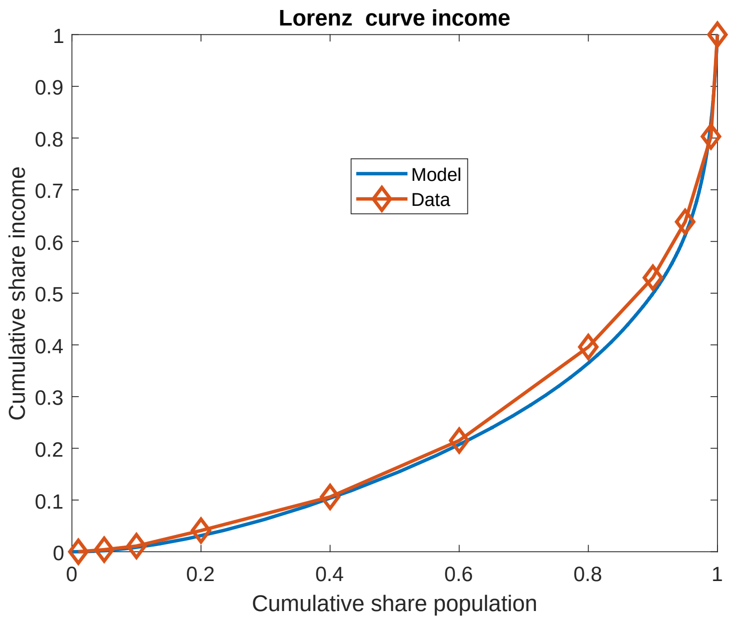


Figure 3: Lorenz curve for income: Calibrated general equilibrium model vs. data. The model with an endogenous income distribution resulting from the human capital accumulation matches the income distribution very well both in the top and bottom end. Data source: SCF 2013 (Kuhn and Riós-Rull, 2016).

learning function namely $q(\ell) = (\ell^{\theta} - 1)/\theta$ with $0 < \theta < 1$. Observe that $q(\ell) = \ln(\ell)$ employed in the previous section is a special nested case for $\theta \to 0$. In similar spirited models Ludwig et al. (2012), Huggett et al. (2011), and Guvenen et al. (2014) set the curvature of human capital production θ to values 0.65, 0.7, and 0.8, respectively. Our setting is robust to this set of values and we chose the intermediate value of 0.75 as the benchmark.

We follow Kindermann and Krueger (2022) for the calibration of the consumption and capital income tax rates. In contrast to the partial equilibrium model, output is not produced by labor only. The presence of alternative tax income sources (capital income and consumption) thus relieves some *fiscal pressure* from the progressive labor income tax system similar to the findings of Heathcote and Tsujiyama (2021). The exogenous government expenses are mainly financed by the income tax with the numbers broadly in line with Kindermann and Krueger (2022) (cf. Table 6). The transfer share matches broadly the time-average of 1.6% observed in data from the Congressional Budget Office for income security programs.

We target and match well standard goals with the calibrated values reported in Table 5. The presence of savings as an alternative for self-insurance requires lower values for

Government budget (model)	
Outlays	
Government expenditures	18.9%
Transfers	1.1%
Tax income	
Labor tax	12.2%
Capital income tax	3.9%
Consumption tax	3.9%

Table 6: Government budget (all measured in percentage of GDP).
The transfer share is broadly consistent with the time-average (1962-2022) of 1.6% observed in data from the Congressional Budget Office for income security programs. These programs include unemployment compensation, Supplemental Security Income, the refundable portion of the earned income and child tax credits, the Supplemental Nutrition Assistance Program, family support, child nutrition, and foster care. Note that the data displays some substantial time variation with peaks in economic downturns such as the aftermath of the Great Financial Crisis or under the recent Covid pandemic 2020/2021.

labor disutility χ relative to the PE to match the average labor supply of $\ell_{max}/3 = 1.4$. The values of depreciation δ_k and TFP are chosen to produce realistic absolute and relative factor prices.

Baseline results: The calibrated model does an excellent job at matching the distribution of labor income captured by the Lorenz curves and displayed in Figure 3. This is in particular noteworthy, since, in contrast to standard models in the tradition of Aiyagari (1994), income is not exogenously imposed but results from an individual optimization in our model. While income is the focus of the model, it also matches key inequality measures of the wealth distribution in excess of income inequality (cf. Table 7).

In contrast to before (with $\gamma \to 1$), individual labor supply is not independent of the level of human capital. There is a tendency of those with higher level of human capital to also work longer hours exacerbating income inequality relative to the PE model. Therefore, we require lower values for the variance of shocks to human capital σ^2 to match the empirically observed Gini-coefficient. In line with Pijoan-Mas (2006) those with higher levels of assets enjoy more leisure and rely to a larger extent on their asset income.[26] It is also important to point out that in contrast to Heathcote et al. (2017) this is not driven by preference heterogeneity for leisure.

Both in the data and in the model the inequality of hours worked is substantially lower than the inequality of income suggesting that there is little variation in hours worked in

[26]Optimal labor supply as a function of financial assets and human capital is displayed in Online Appendix H.

Category	Measure	Data	Model
Income	Gini	0.58	0.58
	Top 10% share	49.0%	50.2%
Wealth	Gini	0.87	0.90
	Top 10% share	74.9%	87.9%
Hours	Gini	0.11	0.06
	Mean/Qi_1	1.4	1.3
	Qi_5/Mean	1.3	1.1

Table 7: Distributional measures for income, wealth and hours worked – Data vs. model. The calibrated model not only matches income inequality, but also inequality of wealth well even slightly overstating some key measures. The measure Qi_j represent hours worked at quintile $j = 1, \cdots, 5$ of the labor supply distribution. Data sources: income and wealth, SCF 2013 (Kuhn and Ríos-Rull, 2016); hours worked, Current Population Survey (CPS) 2002 (Pijoan-Mas, 2006).

the population (cf. Table 7). While matching well the lower labor supply at the bottom end of the distribution, the model slightly underestimates the labor supply at the top end and thereby also underestimates the overall dispersion of hours worked.

Due to the infinite horizon nature of the model, its focus is not to match labor market life-cycle behavior. Yet, we can use the resulting endogenous transition matrices to conduct a forward simulation of individual labor market behavior.[27] When considering the behavior of a young individual (age: 25) that enters the labor market with zero assets ($a_0 = 0$) and with the modal value of human capital h_0, we find that until age 55 labor income (after taxes and transfers) increases by a factor 2.1 with work time roughly constant both broadly in line with US micro evidence (cf. e.g. Wallenius, 2011; Blandin and Peterman, 2019). The higher efficiency wages (due to higher human capital) at higher ages are roughly offset by both lower returns to learning-by-doing (cf. e.g. Imai and Keane, 2004) and higher average tax rates leading to a relatively flat labor supply schedule.

5. Optimal progressive taxes in the full model

Varying tax progressivity: Similar to our analysis with hand-to-mouth agents, we can use this framework to numerically identify an optimal tax progressivity τ_y. Figure 4 presents the changes of different measures of income inequality. Comparable to the partial equilibrium exercise a higher tax progressivity lowers market income inequality

[27]The computational details are explained in Online Appendix G under step 3.

and – of course – inequality after taxes and transfers. Both the overall Gini-coefficient and the more narrow top share measure display similar relative changes. The inequality reduction is comparable to the one observed in the PE model in the same range (cf. Figure 1) albeit a little less pronounced.

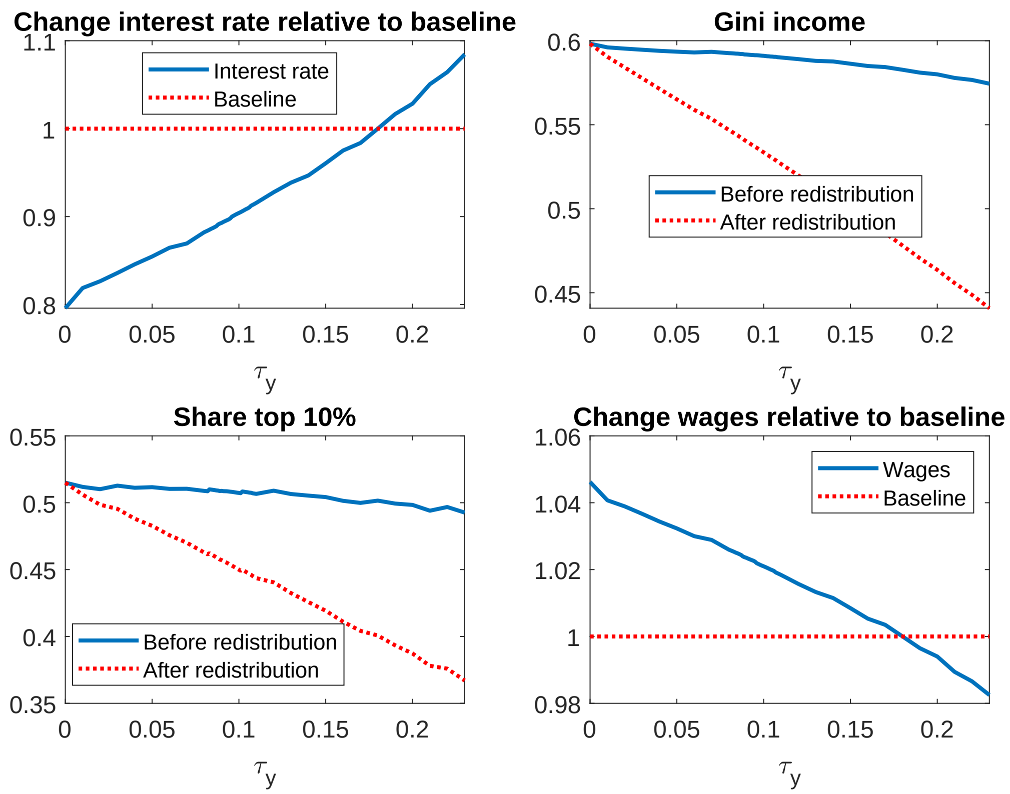
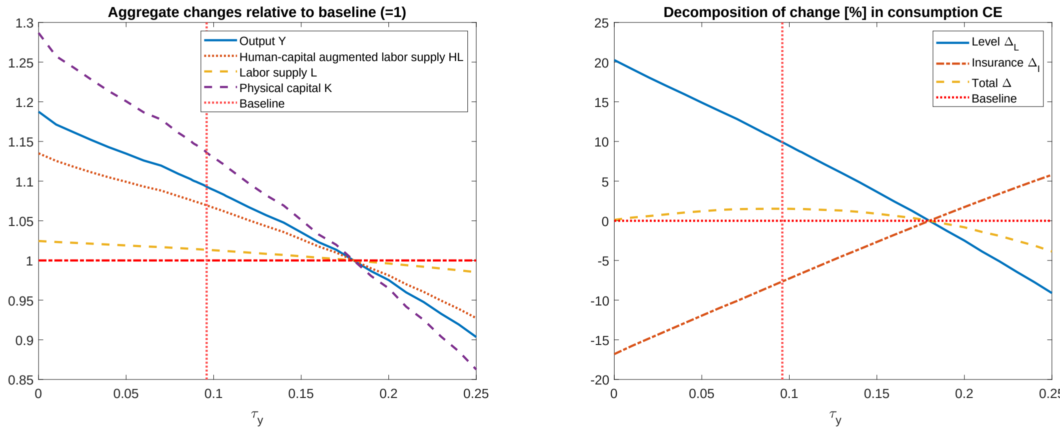


Figure 4: Comparative statics for a variation of tax progressivity τ_y: Income inequality (Gini-coefficient, Top 10% share) and factor prices (interest rate, wages).
Higher tax progressivity lowers inequality both before and after the tax-and-transfer redistribution. Higher tax progressivity changes factor prices by reducing wage rates and increasing interest rates.

The latter is partly due to the fact that in the GE model factor income prices also change. In line with Floden and Lindé (2001), higher taxes require forming less precautionary savings and hence decreases the ratio $\frac{K}{HL}$ manifesting itself in lower wages w and higher interest rates r with the latter being more pronounced. This is an implicit dividend for asset-rich households and confirms that general equilibrium effects often act against the redistributive intentions of the policy maker (Fischer, 2019).

Not surprisingly – and as in our previous analysis with hand-to-mouth agents – higher taxes lower not only the value of immediate labor supply but to an even larger extent the value of aggregate human-capital augmented labor supply (Figure 5(a)). Relative to said previous scenario (cf. Figure 2(a)) the effects are slightly less pronounced since individual labor supply does not react uniformly anymore. In fact, individuals in dire straits with both little human and physical capital will react to the reduction in transfers (following from lower τ_y) by eventually increasing their labor supply.



(a) Aggregate measures

(b) Welfare

Figure 5: The effect of different tax regimes on several aggregate measures (panel a) and the decomposed welfare effect (panel b).
Higher tax progressivity τ_y lowers all aggregate measures most notably capital accumulation and thus average (total) income. For the given parameters the welfare-optimizing taxation is given by $\tau_y^{total} = 9.6\%$.

Moving to a flatter tax scheme increases the dispersion of hours worked. Most importantly, more *efficient* individuals with high human capital are encouraged to work more. Relatively to the partial equilibrium model in which all individuals adjust labor supply identically, there is less of a reaction in aggregate labor supply and human capital and hence in output after a reduction in tax progressivity.

In the general equilibrium economy, individuals can also accumulate physical capital K. Since the progressive tax system lowers the need to self-insure by holding physical assets, the aggregate capital also decreases with higher taxation (Figure 5(a)). This measure shows the highest sensitivity of all input factors to changes in tax progressivity also explaining why interest rates r eventually rise with higher tax progressivity. In other words, while both aggregate human-capital-augmented labor supply HL and physical capital K fall as result of more a progressive income tax scheme, the latter falls by a larger magnitude than the former (Figure 5(a)), leading to rising interest rates (Figure 4 upper left panel) and falling wage levels (Figure 4 lower right panel).

Optimal tax progressivity: Like the scenario with hand-to-mouth agents, in the current situation the social planner balances out aggregate effects on efficiency (level effects Δ_L) with equity concerns (insurance effects Δ_I) as displayed in Figure 5(b). While the level effect is of a similar magnitude like the one observed in the PE model (cf. Figure 2(b)), the insurance effect is substantially lower.

The key difference is that in this extended model, agents are not consuming in a hand-to-mouth fashion, but are forming precautionary savings which help to smooth consumption and to self-insure against adverse shocks to human capital. As such, the insurance property provided by the progressive government-run tax and transfer system is substantially less pronounced relative to the PE model. In fact, the optimal level of tax progressivity balancing off these the level and the insurance effect amounts to $\tau_y = 9.6\%$ associated with a gain in CE consumption of 1.5% very close the baseline result in Heathcote et al. (2017). Not only is this less than the one proposed in the PE exercise, but also less than the status-quo of $\tau_y = 18\%$.

We can decompose the effects of such a policy reform in steady state. We order individuals by total post tax-and-transfer earnings $e(a, h)$ consisting of both labor and capital income – formally $e(a, h) = T(whl(a, h))^{1-\tau_y} + (1 - \tau_K)ra$.

First of all, it is interesting to see that the share of capital income is increasing towards the top end of the earnings distribution (cf. left panel of Figure 6) in line with US evidence from the SCF (Dyrda and Pedroni, 2022). The very top end of the distribution, however, mostly earns labor income. This is in line with detailed evidence from US tax data showing that most of the top income owners derive their income from human capital (Smith et al., 2019). Thus, they are not idle or coupon-clipping rich, but working rich. In the logic of our model, positive shocks to human capital (luck) and their endogenous decision to work long hours endow them with large amounts of human capital. In the model of Brüggemann (2021) the top share of the income distribution are formalized as entrepreneurs, while in Kindermann and Krueger (2022) they are *superstars* with exceptionally high income states. In both cases, earnings at the top are mostly driven by some exogenous stochastic ability not affected by the tax rate.

As shown in the upper half of Table 8, a cut in tax progressivity under the optimal policy on labor income makes all individuals supply more labor. This effect is more pronounced towards the top end of the distribution. The effect is, furthermore, emphasized

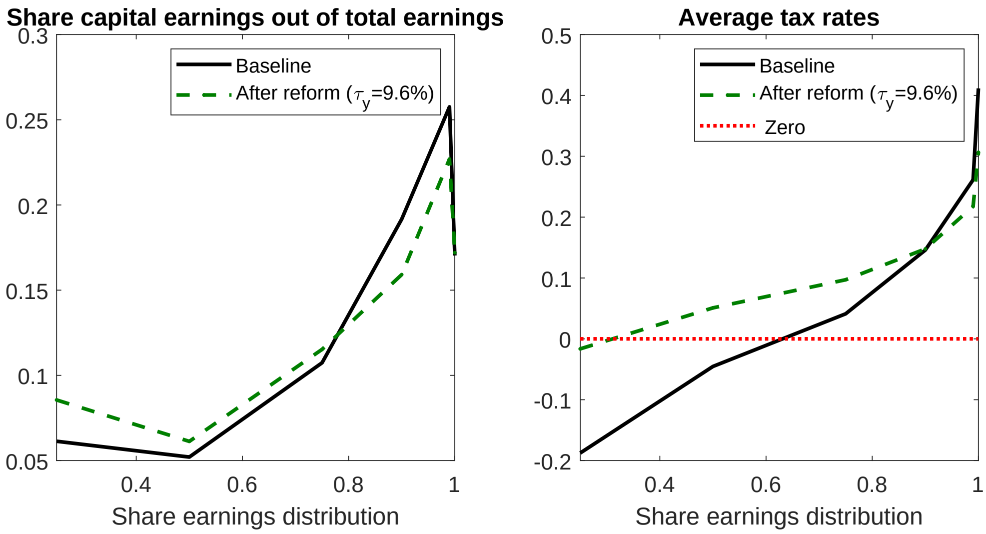


Figure 6: Difference along the earnings distribution. Left panel: Share of capital income out of total earnings in baseline scenario. Right panel: Average tax rate in the baseline and after tax reform ($\tau_y = 9.6\%$).

Earnings e are computed as the sum of labor earnings and capital earnings after taxes and transfers. The share of capital is highest at the top 5%. The very top rely mostly on labor income though. The reduction of tax progressivity flattens the average tax rate. It reduces average taxes for the top 10%, while increasing them for the rest of the population.

by the general equilibrium effect of higher wages. Not only does the reform encourage labor supply, but also the formation of precautionary savings eventually increasing the share of capital earnings out of total earnings at the bottom end of the distribution (cf. left panel of Figure 6). In contrast, individuals at the top end increase their share of labor as their prime source of income after the reduction of average tax rate for the top.

The average tax rates before and after the reform are displayed in the right panel of Figure 6. The winners of a decrease of tax progressivity are at the top end of the distribution. Only a small group of individuals at the top 10% are subject to lower average taxes after the reform. Moreover, the threshold value of being a net transfer receiver also increases following the the reform. This is also emphasized due to the assumption that government expenses are proportional to overall output ($G = g \cdot Y$) also made in Heathcote et al. (2017). One might consider this as reflecting that the overall higher economic activity after a reduction of tax progressivity will also increase consumption of public goods such as streets in the same pace, requiring the government to also increase their expenditures. To maintain a balanced budget the government will cut transfers directed to the bottom end of the earnings distribution.

Scenario	Share	0.0 - 0.25	0.25 - 0.5	0.5 - 0.75	0.75 - 0.9	Top 10%	Top 1%	Aggregate
Endogenous Human	Labor ℓ	0.4	1.0	0.7	2.5	1.2	1.4	1.3
Capital Accumulation	Earnings e	-19.0	-9.2	-6.3	-5.9	1.0	14.9	9.3
$\tau_y^* = 9.6\%$	Consumption c	-11.5	-2.8	0.3	0.1	10.0	21.8	6.7
	CE[c]	-15.8	-4.6	-1.2	-2.6	8.3	19.7	1.5
Exogenous Human	Labor ℓ	-5.9	-6.0	-8.3	-12.0	-17.7	-12.4	-7.9
Capital Accumulation	Earnings e	50.1	14.2	14.4	5.4	-11.3	-33.6	-16.9
$\tau_y^* = 38.0\%$	Consumption c	42.2	10.8	9.9	0.3	-6.0	-30.0	-9.1
	CE[c]	57.9	15.6	15.9	8.3	10.4	-18.9	5.0

Table 8: Decomposition of reaction to switch to optimal tax progressivity with endogenous and exogenous human capital. Percentage change in each bracket of the earnings distribution.

Changes of outcome variables within certain earnings brackets after a change of tax progressivity. With endogenous human capital accumulation with learning-by-doing, the optimal tax progressvity is $\tau_y^* = 9.6\%$. In this case, while labor supply increases along the distribution, gains in terms of earnings after taxes and transfer, consumption, and Certainty Equivalent (CE) consumption fall only to the top end of the distribution. With exogenous human capital the optimal tax progressivity is $\tau_y^* = 38.0\%$ higher than the baseline of 18%. The rise in tax progressivity leads to the basic mirror image reaction with individuals at the top end being worse off after the reform. See also footnote 28 for further explanation on the changes in CE consumption.

Table 8 also displays the changes in consumption within each earnings bracket. For the case with endogenous human capital accumulation, it is only those at the top 25% – and especially the even more narrow share of the top 1% – that get the gain in consumption. The aggregate gains in output and consumption thus obscures from the fact that the gains are highly disproportionally shared along the earnings bracket. Relatively to the status quo such a reform would lead to a rise in of income inequality after tax and transfers by roughly 10% both in terms of Gini and the share of the Top 10% (cf. Figure 4). So, while the reform does improve overall welfare (as measured by Certainty Equivalent Consumption) and hence should be chosen under the veil-of-ignorance, it will unlikely find a democratic majority if individuals assume to remain in their current earnings bracket.[28]

[28] Observe that the the changes in the CE-consumption for the shares are calculated under the assumption of inertia, i.e. assuming individuals imagine themselves remaining part of said group. More formally, this within-group change in CE consumption Δ_p in the percentile p is given by

$$\int_{Q_p'} \mathbb{E}_0 \left[\int_0^\infty \exp(-\rho t) U(c,\ell) dt \right] df(e) \Bigg|_{\tau_y'} = \int_{Q_p^0} \mathbb{E}_0 \left[\int_0^\infty \exp(-\rho t) U([1+\Delta_p]c,\ell) dt \right] df(e) \Bigg|_{\tau_y^0},$$

Social Planner Preference	Endogenous Human Capital Accumulation	Exogenous Human Capital Accumulation
$\tilde{\gamma} = 1$	7%	37.2%
Utilitarian ($\tilde{\gamma} = 1.1$)	9.6%	38%
$\tilde{\gamma} = 1.5$	17%	38.5%
$\tilde{\gamma} = 2$	22%	39.4%

Table 9: Optimal tax progressivity τ_y for different welfare specifications in the GE model
Higher risk aversion $\tilde{\gamma}$ of the social planner warrants higher tax progressivity.

Similar to the PE model, the outcome depends on the redistributive preferences of the social planner as captured by the risk aversion $\tilde{\gamma}$ (cf. eq. 25 and the surrounding discussion). For the baseline, we assume a utilitarian welfare function ($\tilde{\gamma} = \gamma = 1.1$), but – as in the PE model – a rise in $\tilde{\gamma}$ raises the value of tax progressivity perceived as optimal (cf. Table 9). Indeed, for a slightly higher value of $\tilde{\gamma} = 1.5$ – often assumed in these types of model (Brüggemann, 2021; Kindermann and Krueger, 2022) – optimal tax progressivity already rises substantially to a value of $\tau_y = 17\%$ close to the status-quo.

Exogenous human capital: Similar to the partial equilibrium model, we can consider optimal tax progressivity under exogenous human capital.[29] This is a standard model considered thus far in the literature with an exogenous human capital distribution. Both in line with the literature (Brüggemann (2021), Kindermann and Krueger (2022)) and the partial equilibrium exercises, the optimal degree of tax progressivity ($\tau_y = 38\%$) is higher than the prevailing status quo suggesting higher taxes at the top and more transfers at the bottom. For fixed aggregate human capital H aggregate output will react less pronounced to changes in tax progressivity. For example a change to a flat tax system ($\tau_y = 0$) only raises output by 10% as opposed to a substantially larger effect in the presence of endogenous human capital accumulation (cf. Figure 5(a)).

The suggested increases of top tax has basically the mirror image impact to the reduction of tax progressivity proposed under the case with human capital accumulation (cf. Table 8). There is a general reduction in labor supply especially at the top end. The majority of the population experiences rises in earnings after-taxes-and-transfers which are particularly pronounced towards the bottom end. Only the narrow top 1% experience a reduction in CE levels. We observe that higher risk aversion $\tilde{\gamma}$ by the social

where Q_p^0 and Q_p' are the quantiles before and after the tax changes. For the aggregate level it is given by (16) where we use the earnings distribution $f(e)$ instead of that of human capital.

[29]This is under the assumption of no learning ($A = 0$) and an exogenous drift term $m = 2.8\%$ (cf. Table 1). We slightly recalibrate the model to still match the targets (cf. Online Appendix H).

planner warrants higher tax progressivity (cf. Table 9), which is in line with our previous analysis with hand-to-mouth agents. The overall welfare gain of the reform is also higher than under the scenario with endogenous human capital. Given that under exogenous human capital the distribution of earnings is less sensitive to changes in tax progression, optimal tax progressivity is also less sensitive to the choice of $\tilde{\gamma}$ as compared to the case with endogenous human capital accumulation.

Discussion: Finally, what is the optimal level of tax progressivity? This study emphasizes the role of endogenous human capital accumulation that is usually not considered in the literature. Disregarding this channel and thus ignoring long-run effects of discouraging the accumulation of human capital on the job, biases the estimates of optimal levels towards high tax progressivity (cf. e.g. Brüggemann, 2021; Kindermann and Krueger, 2022; Ferrière et al., 2023). Indeed, our finding for the optimal tax progessivity is closely in line with the quantitative result in Heathcote et al. (2017). While their paper features human capital investment, it is considered as a once-for-all decision. This does not capture important dynamic effects relevant in economies with human capital obsolescence requiring individuals to restart their career repeatedly. Their simple model also obscures the fact that a reform that reduces tax progressivity is welfare reducing for the dominant share of the population.

We devote our attention to particular structural form of the tax system (in the framework of Heathcote et al., 2017) and on the tax progressivity τ_y as the single policy instrument. Of course, the policy maker is not restricted to this and welfare improvements are possible by introducing a richer tax system with multiple variables (Ferrière et al., 2023) or even considering a non-parametric approach. Brüggemann (2021) and Kindermann and Krueger (2022) focus on the top end of the distribution and eventually argue that top taxes in the US should be more progressive than the status quo. Both approaches abstract from the long-run impact on human capital.

6. Summary and outlook

This paper develops a model of human capital accumulation through *learning on the job* in the presence of human capital obsolescence shocks. Our PE framework is capable of capturing income inequality both in the bottom and the top ends of the distribution while remaining analytically tractable. With this model we perform a policy analysis to design a progressive tax and transfer system. Optimal taxes trading off equity and efficiency concerns depend crucially on the risk aversion of the social planner. The analysis is then extended to a rich quantitative general equilibrium model in the Bewley-Huggett-Aiyagari tradition in which individuals can self-insure both by adjusting labor supply and holding financial assets. Relative to models with exogenous income, the presence of endogenous human capital accumulation calls for substantially lower tax progressivity when self-insurance through precautionary savings is an option. By ignoring this channel, standard models with exogenous income tend to be biased towards high level of tax progressivity. Our rich model, however, also reveals that after a reform that reduces tax progressivity, welfare effects are very heterogeneous along the earnings distribution. The winners of such policy in terms of certainty equivalent consumption are the top 10% share of the income distribution, while the losers are the bottom 90%.

A potential next step worth exploring is a complete system of optimal taxes including taxes on capital income. The growing literature has thus far explored this question in models with exogenous human capital (cf. e.g. Boar and Midrigan, 2022; Carroll et al., 2023) and thus might produce biased estimates. Relatedly, Blandin and Peterman (2019) argue that the presence of a learning-by-doing channel warrants a lower tax progressivity on labor income and a higher (flat) tax on capital income relative to a model with purely exogenous income. In their model focusing on life-cycle features rather than cross-sectional inequality, this is especially important for not discouraging labor supply of older individuals rich in both physical and human capital associated with high marginal tax brackets. In this paper, we isolated the impact of progressive labor income taxation. Optimal tax progressivity taxing labor and capital income jointly or alternatively separately, could thus be a topic for future work. Given that taxes and transfers are decided in a democratic process, it might also be worthwhile to explore the difference between normative recommendations from theoretical models and the empirically observed tax systems through a political economic lens. We leave these avenues to future research.

Acknowledgments

We are indebted to Tobias Broer, Emiliano Santoro, and Alexis Akira Toda for detailed comments. We would like to thank Philipp Grübener, Karl Harmenberg, Benjamin Larin, William Peterman, and Kjetil Storesletten for helpful comments, and, more generally, participants at the CEF 2023, the Oslo Macro Group, VfS 2021, Jönköping University, and at the Copenhagen Macro seminar.

The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) and later National Academic Infrastructure for Supercomputing in Sweden (NAISS) at the Centre for Scientific and Technical Computing at Lund University (LUNARC – with project number: LU 2021/2-68), and later at the National Supercomputer Center (NSC) partially funded by the Swedish Research Council through grant agreement no. 2018-05973 (project numbers: SNIC 2022/22-195 and later NAISS 2023/5-100).

Danial Ali Akbari gratefully acknowledges for financial support from the Thule Foundation through the Skandia Research Program, Jan Wallander and Tom Hedelius Foundation, Siamon Foundation and the Foundation for Economic Research at Lund University. Thomas Fischer gratefully acknowledges financial support by the Jan Wallander and Tom Hedelius Foundation and FORMAS.

References

Acemoglu, D., G. Gancia, and F. Zilibotti (2012). Competing engines of growth: Innovation and standardization. *Journal of Economic Theory 147*(2), 570–601.3.

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2022). Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach. *Review of Economic Studies 89*(1), 45–86.

Aiyagari, S. R. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics 109*(3), 659–84.

Aoki, S. and M. Nirei (2017). Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the United States. *American Economic Journal: Macroeconomics 9*(3), 36–71.

Autor, D. H., D. Dorn, and G. H. Hanson (2016). The China shock: Learning from labor-market adjustment to large changes in trade. *Annual Review of Economics 8*, 205–240.

Autor, D. H. and M. J. Handel (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labor Economics 31*(S1), S59–S96.

Badel, A., M. Huggett, and W. Luo (2020). Taxing Top Earners: a Human Capital Perspective. *Economic Journal 130*(629), 1200–1225.

Bakis, O., B. Kaymak, and M. Poschke (2015). Transitional Dynamics and the Optimal Progressivity of Income Redistribution. *Review of Economic Dynamics 18*(3), 679–693.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy 75*, 352–352.

Blanchard, O. J. (1985). Debt, Deficits, and Finite Horizons. *Journal of Political Economy 93*(2), 223–47.

Blandin, A. (2018). Learning by doing and ben-porath: Life-cycle predictions and policy implications. *Journal of Economic Dynamics and Control 90*, 220–235.

Blandin, A. and W. B. Peterman (2019). Taxing capital? The importance of how human capital is accumulated. *European Economic Review 119*(C), 482–508.

Boar, C. and V. Midrigan (2022). Efficient redistribution. *Journal of Monetary Economics 131*(C), 78–91.

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

Bénabou, R. (1996). Inequality and Growth. In *NBER Macroeconomics Annual 1996, Volume 11*, NBER Chapters, pp. 11–92. National Bureau of Economic Research, Inc.

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

Carroll, D. R., A. Luduvice, and E. Young (2023). Optimal Fiscal Reform with Many Taxes. Working Papers 23-07, Federal Reserve Bank of Cleveland.

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014). Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility. *American Economic Review 104*(5), 141–47.

Cocco, J. F., F. J. Gomes, and P. J. Maenhout (2005). Consumption and Portfolio Choice over the Life Cycle. *The Review of Financial Studies 18*(2), 491–533.

Coffey, T., C. Kelley, and D. Keyes (2003). Pseudotransient continuation and differential-algebraic equations. *SIAM Journal on Scientific Computing 25*, 553–569.

Diamond, P. and E. Saez (2011). The Case for a Progressive Tax: From Basic Research to Policy Recommendations. *Journal of Economic Perspectives 25*(4), 165–90.

Dyrda, S. and M. Pedroni (2022). Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk. *The Review of Economic Studies 90*(2), 744–780.

Ebenstein, A., A. Harrison, M. McMillan, and S. Phillips (2014). Estimating the impact of trade and offshoring on american workers using the current population surveys. *The Review of Economics and Statistic 96*(4), 581–595.

Ferrière, A., P. Grübener, G. Navarro, and O. Vardishvili (2023). On the optimal design of transfers and income-tax progressivity. *Journal of Political Economy - Macroeconomics 1*(2), 276–333.

Fischer, T. (2019). Determinants of Wealth Inequality and Mobility in General Equilibrium. Working Papers 2019:22, Lund University, Department of Economics.

Floden, M. and J. Lindé (2001). Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance? *Review of Economic Dynamics 4*(2), 406–437.

Flodén, M. (2001). The effectiveness of government debt and transfers as insurance. *Journal of Monetary Economics 48*(1), 81–108.

French, E. (2005). The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour. *Review of Economic Studies 72*(2), 395–427.

Gabaix, X. (2009). Power Laws in Economics and Finance. *Annual Review of Economics 1*(1), 255–294.

Gabaix, X., J. Lasry, P. Lions, and B. Moll (2016). The Dynamics of Inequality. *Econometrica 84*, 2071–2111.

Gathmann, C. and U. Schönberg (2010). How general is human capital? A task-based approach. *Journal of Labor Economics 28*(1), 1–49.

Gomez, M. (2015). Solving PDEs Associated with Economic Models. Technical report.

Goos, M., A. Manning, and A. Salomons (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review 104*(8), 2509–26.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2021). What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics? *Econometrica 89*(5), 2303–2339.

Guvenen, F., B. Kuruscu, and S. Ozkan (2014). Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis. *Review of Economic Studies 81*(2), 818–850.

Heathcote, J., F. Perri, G. Violante, and L. Zhang (2023). More Unequal We Stand? Inequality Dynamics in the United States, 1967–2021. *Review of Economic Dynamics 50*, 235–266.

Heathcote, J., K. Storesletten, and G. L. Violante (2017). Optimal Tax Progressivity: An Analytical Framework. *The Quarterly Journal of Economics 132*(4), 1693–1754.

Heathcote, J., K. Storesletten, and G. L. Violante (2020a). Optimal progressivity with age-dependent taxation. *Journal of Public Economics 189*, 104074.

Heathcote, J., K. Storesletten, and G. L. Violante (2020b). Presidential Address 2019: How Should Tax Progressivity Respond to Rising Income Inequality? *Journal of the European Economic Association 18*(6), 2715–2754.

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

Huggett, M., G. Ventura, and A. Yaron (2011). Sources of lifetime inequality. *American Economic Review 101*(7), 2923–54.

Hémous, D. and M. Olsen (2022). The Rise of the Machines: Automation, Horizontal Innovation, and Income Inequality. *American Economic Journal: Macroeconomics 14*(1), 179–223.

Imai, S. and M. P. Keane (2004). Intertemporal Labor Supply and Human Capital Accumulation. *International Economic Review 45*(2), 601–641.

Jedwab, R., P. Romer, A. M. Islam, and R. Samaniego (2023). Human capital accumulation at work: Estimates for the world and implications for development. *American Economic Journal: Macroeconomics 15*(3), 191–223.

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

Kapička, M. (2015). Optimal Mirrleesean Taxation in a Ben-Porath Economy. *American Economic Journal: Macroeconomics 7*(2), 219–248.

Kapička, M. and J. Neira (2019). Optimal Taxation with Risky Human Capital. *American Economic Journal: Macroeconomics 11*(4), 271–309.

Kindermann, F. and D. Krueger (2022). High marginal tax rates on the top 1 percent? lessons from a life-cycle model with idiosyncratic income risk. *American Economic Journal: Macroeconomics 14*(2), 319–66.

Kredler, M. (2014). Experience vs. obsolescence: A vintage-human-capital model. *Journal of Economic Theory 150*, 709–739.

Krueger, D., K. Mitman, and F. Perri (2016). *Macroeconomics and Household Heterogeneity*, Volume 2 of *Handbook of Macroeconomics*, Chapter 0, pp. 843–921. Elsevier.

Kuhn, M. and J.-V. Riós-Rull (2016). 2013 Update on the U.S. Earnings, Income, and Wealth Distributional Facts: A View from Macroeconomics. *Quarterly Review* (April), 1–75.

Lazear, E. P. (2009). Firm-specific human capital: A skill-weights approach. *Journal of Political Economy 117*(5), 914–940.

Ludwig, A., T. Schelkle, and E. Vogel (2012). Demographic change, human capital and welfare. *Review of Economic Dynamics 15*(1), 94–107.

Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of Labor Economics 13*(4), 653–677.

Nuño, G. and B. Moll (2018). Social optima in economies with heterogeneous agents. *Review of Economic Dynamics 28*, 150–180.

Peterman, W. (2016). The effect of endogenous human capital accumulation on optimal taxation. *Review of Economic Dynamics 21*, 46–71.

Pijoan-Mas, J. (2006). Precautionary Savings or Working Longer Hours? *Review of Economic Dynamics 9*(2), 326–352.

Piketty, T. (2014). *Capital in the twenty-first century.* Boston: Harvard University Press.

Piketty, T. and E. Saez (2007). How Progressive is the U.S. Federal Tax System? A Historical and International Perspective. *Journal of Economic Perspectives 21*(1), 3–24.

Piketty, T. and E. Saez (2013). Chapter 7 - optimal labor income taxation. In A. J. Auerbach, R. Chetty, M. Feldstein, and E. Saez (Eds.), *handbook of public economics, vol. 5*, Volume 5 of *Handbook of Public Economics*, pp. 391–474. Elsevier.

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

Saez, E. and S. Stantcheva (2018). A simpler theory of optimal capital taxation. *Journal of Public Economics 162*, 120–142.

Smith, M., D. Yagan, O. Zidar, and E. Zwick (2019). Capitalists in the Twenty-First Century. *The Quarterly Journal of Economics 134*(4), 1675–1745.

Stantcheva, S. (2017). Optimal Taxation and Human Capital Policies over the Life Cycle. *Journal of Political Economy 125*(6), 1931–1990.

Straub, L. and I. Werning (2020). Positive Long-Run Capital Taxation: Chamley-Judd Revisited. *American Economic Review 110*(1), 86–119.

Sullivan, P. (2010). Empirical evidence on occupation and industry specific human capital. *Labour Economics 17*(3), 567–580.

Toda, A. A. (2012). The double power law in income distribution: Explanations and evidence. *Journal of Economic Behavior & Organization 84*(1), 364 – 381.

Toda, A. A. and K. Walsh (2015). The Double Power Law in Consumption and Implications for Testing Euler Equations. *Journal of Political Economy 123*(5), 1177–1200.

Wallenius, J. (2011). Human Capital Accumulation and the Intertemporal Elasticity of Substitution of Labor: How Large is the Bias? *Review of Economic Dynamics 14*(4), 577–591.

Wu, C. (2021). More unequal income but less progressive taxation. *Journal of Monetary Economics 117*(C), 949–968.

A. Optimal human capital accumulation

The utility function of individuals is given by:

$$U(c,\ell) = \ln(c) - \chi\frac{\ell^{1+\frac{1}{\upsilon}}}{1+\frac{1}{\upsilon}}, \tag{34}$$

and they face the following constraint:

$$c = T(wh\ell)^{1-\tau_y}. \tag{35}$$

Combining both we find a modified utility only depending on labor $U(\ell)$ being the only available control in the HJB:

$$U(\ell) = \ln(T) + (1-\tau_y)[\ln(h) + \ln(w)] + (1-\tau_y)\ln(\ell) - \chi\frac{\upsilon}{1+\upsilon}\ell^{\frac{1+\upsilon}{\upsilon}}. \tag{36}$$

Now, we can solve the HJB of the dynamic problem:

$$\rho V = \max_{\ell}\left\{U(\ell) + V_h(A\ln(\ell)h - \delta_h h) + V_{hh}\frac{1}{2}\sigma^2 h^2 + p(V(h_0) - V)\right\}. \tag{37}$$

We guess a value function of the type $V(h) = K_2 + K_1\ln(h)$ depending on the state h and some constants K_1 and K_2 to be determined. The first-order condition with respect to labor supply ℓ is given by:

$$V_h\frac{A}{\ell}h = \chi\ell^{\frac{1}{\upsilon}} - \frac{1-\tau_y}{\ell}. \tag{38}$$

Using the value function we can therefore derive:

$$\ell = \left(\frac{AK_1 + (1-\tau_y)}{\chi}\right)^{\frac{\upsilon}{1+\upsilon}}. \tag{39}$$

Note that the optimal labor supply is independent of human capital h. This is due to the log-utility function in consumption also implying that the value function of the state h will be of the log-type.

Inserting equation 39 and 36 jointly with the value function into the HJB leads to:

$$(\rho+p)(K_2+K_1\ln(h)) = \ln(T)+(1-\tau_y)\ln(h)+\frac{\upsilon}{1+\upsilon}\ln\left(\frac{(1-\tau_y)+AK_1}{\chi}\right)(1-\tau_y)$$
$$+(1-\tau_y)\ln(w)-\frac{1+\upsilon}{\upsilon}[(1-\tau_y)+AK_1]+K_1\left[A\frac{\upsilon}{1+\upsilon}\ln\left(\frac{(1-\tau_y)+AK_1}{\chi}\right)-\delta_h-\frac{1}{2}\sigma^2\right]+pK_2. \quad (40)$$

Note that $h_0=1$ implies $V(h_0)=K_2$. By comparing coefficients related to $\ln(h)$ it is easy to see that $K_1=\frac{1-\tau_y}{\rho+p}$ [30], thus implying an optimal labor supply of:

$$\ell = \left(\frac{(1-\tau_y)(A+\rho+p)}{\chi(\rho+p)}\right)^{\frac{\upsilon}{1+\upsilon}}. \quad (41)$$

Using equation 36 jointly with equation 41 leads to optimal consumption which (in the absence of savings) is equal to labor income:

$$y \sim c = T(wh)^{1-\tau_y}\left(\frac{(1-\tau_y)(A+\rho+p)}{\chi(\rho+p)}\right)^{\frac{(1-\tau_y)\upsilon}{1+\upsilon}}. \quad (42)$$

It is convenient to rewrite it in logs allowing us to interpret the respective elasticities:

$$\ln(y)\sim\ln(T)+(1-\tau_y)[\ln(h)+\ln(w)]+\frac{(1-\tau_y)\upsilon}{1+\upsilon}\ln\left(\frac{(1-\tau_y)(A+\rho+p)}{\chi(\rho+p)}\right). \quad (43)$$

The latter is also identical to utility out of consumption.

Inserting the result into the general equation leads to the following drift parameter:

$$m = A\frac{\upsilon}{1+\upsilon}\ln\left(\frac{(1-\tau_y)(A+\rho+p)}{\chi(\rho+p)}\right)-\delta_h. \quad (44)$$

B. Stationary distribution as the solution of the Kolmogorov-Forward equation

The Kolmogorov-Forward (or alternatively Fokker-Planck) equation describing the time-evolution of the human capital distribution reads:

$$\frac{\partial f(h,t)}{\partial t} = -\frac{\partial}{\partial X}[m\cdot h\cdot f(h,t)]+\frac{\partial^2}{\partial h^2}\left[\frac{1}{2}\sigma^2h^2\cdot f(h,t)\right]-pf(h,t)+p\hat{\delta}(h-h_0), \quad (45)$$

for which $\hat{\delta}$ is the Dirac delta.

[30] For the sake of completeness the second free parameter is given by: $K_2 = \frac{1}{\rho}\left(\ln(T)+(1-\tau_y)\ln(w)+\left(\frac{(1-\tau_y)(A+\rho+p)}{(\rho+p)}\right)\frac{\upsilon}{1+\upsilon}\ln\left(\frac{(1-\tau_y)(A+\rho+p)}{(\rho+p)\chi}\right)-\frac{1-\tau_y}{\rho+p}(\delta_h+\frac{1}{2}\sigma^2)-\frac{1+\upsilon}{\upsilon}\frac{(1-\tau_y)(A+\rho+p)}{(\rho+p)}\right)$.

Following Gabaix (2009), we guess and verify the power-law $f(h) = \hat{C}h^{-\vartheta-1}$ as the stationary solution ($\frac{\partial f(h,t)}{\partial t} \stackrel{!}{=} 0$) by inserting it into the Fokker-Planck-equation:

$$\begin{aligned}\frac{\partial f(h,t)}{\partial t} \stackrel{!}{=} 0 = -\frac{\partial}{\partial h}[(m \cdot h) \cdot \hat{C}h^{-\vartheta-1}] + \frac{\partial^2}{\partial h^2}[\frac{1}{2}\sigma^2 h^2 \cdot \hat{C}h^{-\vartheta-1}] - p\hat{C}h^{-\vartheta-1} \\ = \hat{C}h^{-\vartheta-1}\left[m\vartheta - \frac{1}{2}\sigma^2\vartheta(1-\vartheta) - p\right] \stackrel{!}{=} 0.\end{aligned} \tag{46}$$

The characteristic equation is quadratic and solved by:

$$\begin{aligned}\vartheta^2 + \frac{m - \frac{1}{2}\sigma^2}{\frac{1}{2}\sigma^2}\vartheta - \frac{2p}{\sigma^2} = a^2 - 2A_1 a - A_2 \stackrel{!}{=} 0 \\ \vartheta_{1,2} = A_1 \pm \sqrt{A_1^2 + A_2},\end{aligned} \tag{47}$$

with $A_1 = -\frac{m - \frac{1}{2}\sigma^2}{\sigma^2}$ and $A_2 = \frac{2p}{\sigma^2}$.

C. Inequality measures for the Double Pareto distribution

This reports closed expression for all distributional measures reported in Table 1.

$$\text{Var}(\ln(X)) = \frac{\vartheta_1^2 + \vartheta_2^2}{(\vartheta_1\vartheta_2)^2}. \tag{48}$$

$$\text{Gini}(X) = \frac{2\vartheta_1^2 - 2\vartheta_1\vartheta_2 + 2\vartheta_2^2 - \vartheta_1 - \vartheta_2}{(\vartheta_1 - \vartheta_2)(2\vartheta_1 - 1)(1 - 2\vartheta_2)}. \tag{49}$$

Under negative skewness,[31] in line with the data, we have:

$$Q_{90}/Q_{50} = \left(\frac{\vartheta_1}{\hat{C}}0.1\right)^{-\frac{1}{\vartheta_1}} \left(-\frac{\vartheta_2}{\hat{C}}0.5\right)^{\frac{1}{\vartheta_2}} \tag{50}$$

$$Q_{50}/Q_{10} = \left(\frac{0.5}{0.1}\right)^{-\frac{1}{\vartheta_2}} \tag{51}$$

The mean-median ratio is:

$$E(X)/Q_{50} = \frac{\vartheta_1\vartheta_2}{(1-\vartheta_1)(1-\vartheta_2)} \left(-\frac{\vartheta_2}{\hat{C}}0.5\right)^{\frac{1}{\vartheta_2}} \tag{52}$$

[31] This implies that the median is smaller than the mode X_0 as well as the median and the 10 percentile are on the same part of the distribution

The top-shares $z > \frac{\vartheta_1}{\vartheta_1 - \vartheta_2}$ are given by:

$$s_z = z^{1-1/\vartheta_1} \left(\frac{-\vartheta_2}{\vartheta_1 - \vartheta_2} \right)^{\frac{1}{\vartheta_1}} \frac{\vartheta_2 - 1}{\vartheta_2}. \tag{53}$$

This follows from using the Lorenz-curve $L(F)$ which we can compute from the cumulative density function F. The share of the top $z\%$ is defined as $s_z = 1 - L(1 - z)$.

D. Welfare decomposition in the PE model

Welfare function: The overall utilitarian welfare function in the steady state is given by:

$$\begin{aligned}
W(\tau_y) &= \int U(c,\ell) df(h) \Big|_{\tau_y} \\
&= (1-\tau_y)\left[\ln(\ell(\tau_y)) - \frac{\upsilon}{1+\upsilon}\frac{A+\rho+p}{\rho+p}\right] + \ln(T(\tau_y)) + W_{ins,ut}(\tau_y) \\
&= (1-\tau_y)\left[\ln(\ell(\tau_y)) - \frac{\upsilon}{1+\upsilon}\frac{A+\rho+p}{\rho+p}\right] \\
&\quad + \underbrace{\ln(1-g) + \ln(h_0) + \tau_y \ln(\ell(\tau_y)) + \ln\left(\frac{(1-\vartheta_1(\tau_y)-\tau_y)(1-\vartheta_2(\tau_y)-\tau_y)}{(1-\vartheta_1(\tau_y))(1-\vartheta_2(\tau_y))}\right)}_{\ln(T(\tau_y))} \\
&\quad + \underbrace{(1-\tau_y)\frac{\vartheta_1(\tau_y)+\vartheta_2(\tau_y)}{\vartheta_1(\tau_y)\vartheta_2(\tau_y)}}_{W_{ins,ut}(\tau_y)=(1-\tau_y)\int \ln(h)df(h)=\ln(\mathrm{CE}[h^{1-\tau_y}])} \\
&= \underbrace{\frac{\upsilon}{1+\upsilon}\ln\left(\frac{(1-\tau_y)(A+\rho+p)}{\chi(\rho+p)}\right) - \frac{\upsilon}{1+\upsilon}(1-\tau_y)\frac{A+\rho+p}{\rho+p}}_{U(H(\tau_y)\cdot\ell(\tau_y),\ell(\tau_y)) - \ln(H(\tau_y)) = \ln(\ell(\tau_y)) + v(\ell(\tau_y))} \\
&\quad + \underbrace{\ln\left(\frac{(1-\vartheta_1(\tau_y)-\tau_y)(1-\vartheta_2(\tau_y)-\tau_y)}{(1-\vartheta_1(\tau_y))(1-\vartheta_2(\tau_y))}\right)}_{\ln(h_{TF}^{\tau_y})} + (1-\tau_y)\frac{\vartheta_1(\tau_y)+\vartheta_2(\tau_y)}{\vartheta_1(\tau_y)\vartheta_2(\tau_y)} + \ln(1-g)
\end{aligned} \tag{54}$$

which simplifies due to $h_0 = 1$ implying $\ln(h_0) = 0$. The different components can be neatly described in a additive manner. It is interesting to point out that the level of g just shifts the level of welfare for all τ_y and hence will not impact on the optimal tax rate.

The welfare related to transfers follows from equation 14 and computes as follows:

$$\begin{aligned}
\ln(T) &= \tau_y(\ln(h_{TF} + \ln(\ell(\tau_y))) = \tau_y \ln(\ell(\tau_y)) + \ln(1-g) \\
&+ \ln[(1-\vartheta_1(\tau_y)-\tau_y)(1-\vartheta_2(\tau_y)-\tau_y)] - \ln[(1-\vartheta_1(\tau_y))(1-\vartheta_2(\tau_y))].
\end{aligned} \tag{55}$$

For given values of ϑ_1 and ϑ_2 independent of current taxation τ_y, this is a concave function in τ_y. Yet, as τ_y changes so does the drift parameter of the human capital accumulation process. In general an increase in taxation τ_y lowers m and subsequently increases the value of both Pareto tails. This not only implies less top inequality (before redistribution), but also a lower mean income available for redistribution.

As presented in Proposition 2.2 for the prevailing process the distribution is described by a double Pareto distribution. The latter combines conveniently with the assumed log-utility to compute the measure of consumption insurance measure in the utilitarian case:

$$W_{ins,ut} = \ln(\text{CE}[h^{1-\tau_y}]) = (1-\tau_y)\left(\int_0^{h_0} \ln(h)df(h) + \int_{h_0}^{\infty} \ln(h)df(h)\right), \tag{56}$$

which for the assumed case of $h_0 = 1$ implies:

$$\begin{aligned}(1-\tau_y)\frac{-\vartheta_1(\tau_y)\vartheta_2(\tau_y)}{\vartheta_1(\tau_y)-\vartheta_2(\tau_y)}\left[\int_0^1 \ln(h)h^{-(\vartheta_2(\tau_y)+1)}dh + \int_1^{\infty} \ln(h)h^{-(\vartheta_1(\tau_y)+1)}dh\right]& \\ = \frac{\vartheta_1(\tau_y)+\vartheta_2(\tau_y)}{\vartheta_1(\tau_y)\vartheta_2(\tau_y)}(1-\tau_y).&\end{aligned} \tag{57}$$

Changes in CE: Starting from equation 54 and equation 17 we find:

$$1+\Delta = \exp(v(\tau_y') - v(\tau_y^0)) \cdot \frac{\ell(\tau_y')}{\ell(\tau_y^0)} \cdot \frac{h_{TF}^{\tau_y'}(\tau_y')}{h_{TF}^{\tau_y^0}(\tau_y^0)} \cdot \frac{\text{CE}[h^{1-\tau_y'}(\tau_y')]}{\text{CE}[h^{1-\tau_y'}(\tau_y')]}. \tag{58}$$

The pure level effect Δ_L follows from:

$$\begin{aligned}\ln((1+\Delta_L)C(\tau_y^0)) + v(\ell(\tau_y^0)) &= \ln(C(\tau_y')) + v(\ell(\tau_y')) \\ \rightarrow 1+\Delta_L &= \frac{C(\tau_y')}{C(\tau_y^0)} \cdot \exp\left(v(\ell(\tau_y')) - v(\ell(\tau_y^0))\right) \\ &= \frac{H(\tau_y')}{H(\tau_y^0)} \cdot \frac{\ell(\tau_y')}{\ell(\tau_y^0)} \cdot \exp\left(v(\ell(\tau_y')) - v(\ell(\tau_y^0))\right).\end{aligned} \tag{59}$$

From both equations 58 and 59 and the decomposition (equation 19) follows then:

$$1+\Delta_I = \frac{h_{TF}^{\tau_y'}(\tau_y')}{h_{TF}^{\tau_y^0}(\tau_y^0)} \cdot \frac{\text{CE}[h^{1-\tau_y'}(\tau_y')]/H(\tau_y')}{\text{CE}[h^{1-\tau_y^0}(\tau_y^0)]/H(\tau_y^0)}. \tag{60}$$

Exogenous human capital: For the case of exogenous Pareto tails independent of τ_y with $A = 0$, we have for the level part Δ_L:

$$1+\Delta_L = \exp\left(\frac{\nu}{1+\nu}(\tau_y' - \tau_y^0)\right) \cdot \left(\frac{1-\tau_y'}{1-\tau_y^0}\right)^{\frac{\nu}{1+\nu}}. \tag{61}$$

Optimizing this term is equivalent to optimizing $\ln((1+\Delta_L)(1+\Delta_L(\tau_y^0)) = v(\ell(\tau_y)) + \ln(\ell(\tau_y))$ which gives rise to the following first-order condition:

$$\frac{d(v(\ell(\tau_y)) + \ln(\ell(\tau_y))}{d(1-\tau_y)} = \frac{\nu}{1+\nu}\left(-1 + \frac{1}{1-\tau_y}\right) = 0, \tag{62}$$

which is satisfied for $\tau_y = 0$.

For the insurance part, we have:

$$1 + \Delta_I = \exp\left(\frac{\vartheta_1 + \vartheta_2}{\vartheta_1\vartheta_2}\frac{1-\tau_y'}{1-\tau_y^0}\right)\frac{(1-\tau_y'-\vartheta_1)(1-\tau_y'-\vartheta_2)}{(1-\tau_y^0-\vartheta_1)(1-\tau_y^0-\vartheta_2)}. \tag{63}$$

Optimizing this term is equivalent to optimizing $\ln((1+\Delta_I)(1+\Delta_I(\tau_y^0)) = W_{ins,ut} + \ln(h_{TF}^{\tau_y})$ which gives rise to the following first-order condition:

$$\frac{d(W_{ins,ut} + \ln(h_{TF}^{\tau_y}))}{d(1-\tau_y)} = \frac{\vartheta_1+\vartheta_2}{\vartheta_1\vartheta_2} + \frac{1}{1-\tau_y-\vartheta_1} + \frac{1}{1-\tau_y-\vartheta_2} = 0. \tag{64}$$

This is satisfied by $\tau_y = 1$.[32] This is also the optimal solution for the case of $\nu = 0$ (inelastic labor supply) for there will be no level effect (captured in equation 61 and hence the distribution (captured by the Pareto tails ϑ_1 and ϑ_2) will also be invariant to changes in the tax progressivity τ_y.

Alternative welfare aggregators: For the case of an alternative welfare aggregator we have:

$$\ln(\mathrm{CE}_{\hat\gamma}[c])(\tau_y) = \tau_y \ln(h_{TF}) + \ln(\ell(\tau_y)) + \underbrace{\frac{1}{1-\hat\gamma}\ln\left(\int_0^\infty h^{(1-\hat\gamma)(1-\tau_y)} df(h)\right)}_{W_{ins,\hat\gamma}(\tau_y)}. \tag{65}$$

Using the PDF of the double Pareto distribution (cf. Proposition 2.2 under $h_0 = 1$) we can derive:

$$W_{ins,\hat\gamma} = \frac{1}{1-\hat\gamma}\ln\left(\frac{\vartheta_1(\tau_y)\vartheta_2(\tau_y)}{((1-\tau_y)(1-\tilde\gamma)-\vartheta_1(\tau_y))((1-\tau_y)(1-\tilde\gamma)-\vartheta_2(\tau_y))}\right). \tag{66}$$

We can verify that for $\lim_{\hat\gamma\to 1} W_{ins,\hat\gamma}(\tau_y) = W_{ins,ut}(\tau_y)$ and hence $W_{\tilde\gamma=1}(\tau_y) = W(\tau_y)$.

[32] Alternatively, there is also a highly regressive solution $1-\tau_y = \frac{\vartheta_1^2+\vartheta_2^2}{\vartheta_1+\vartheta_2}$ which for the given calibration does not imply an ergodic distribution and hence can be excluded.

Description	$\upsilon = 0.4$	$\upsilon = 0.2$	$\upsilon = 1$
Labor disutility χ	0.526	0.227	0.872
Optimal taxes τ_y			
Utilitarian ($\tilde{\gamma} = 1$)	0.28	0.39	0.17
+Government ($\phi > 0$)	0.20	0.32	0.10
$\tilde{\gamma} = 0$	0.00	0.00	0.00
$\tilde{\gamma} = 2$	0.50	0.57	0.42
$\tilde{\gamma} \to \infty$	1.00	1.00	1.00

Table 10: Robustness check with a variation of Frisch elasticity υ

For the case of $\tilde{\gamma} = 0$ we have:

$$W_{\tilde{\gamma}=0}(\tau_y) = U(H \cdot \ell, \ell) + \ln(1-g) - \ln(H) + \ln\left(\frac{(1-\vartheta_1(\tau_y)-\tau_y)(1-\vartheta_2(\tau_y)-\tau_y)}{(1-\vartheta_1(\tau_y))(1-\vartheta_2(\tau_y))}\right)$$
$$+ \underbrace{\ln\left(\frac{\vartheta_1(\tau_y)\vartheta_2(\tau_y)}{(1-\tau_y-\vartheta_1(\tau_y))(1-\tau_y-\vartheta_2(\tau_y))}\right)}_{W_{ins,\tilde{\gamma}=0}} = U(H \cdot \ell, \ell) + \ln(1-g), \tag{67}$$

since $\ln(H(\tau_y)) = \ln\left(\frac{\vartheta_1(\tau_y)\vartheta_2(\tau_y)}{(\vartheta_1(\tau_y)-1)(\vartheta_2(\tau_y)-1)}\right)$.

Sensitivity to Frisch elasticity: Compared to our benchmark value $\upsilon = 0.4$ we vary the Frisch elasticity, which also requires reparametrizing χ in order to match the current state. Higher Frisch elasticities associated with a more pronounced reaction to tax changes require larger values of labor disutility χ to match $\ell = 1.4$. As shown in Table 10 an increase of the Frisch elasticity is accompanied by substantial lower optimal taxes and vice versa for a given welfare specification. It is interesting to point out that for a large $\upsilon = 1$, the status quo is almost optimal according to a standard welfare criterion. Indeed, this specification implies that a slight tax reduction would actually (slightly) increase welfare.

E. Taxation in the transition path

So far we only considered steady-state and thereby long-run outcomes. Changes in tax progressivity, will impact individual labor promptly. The impact on overall human capital and on its distribution is, however, more gradual. Several studies have documented that when taking the transition into account in the welfare analysis, higher taxes are warranted (cf. e.g. Bakis et al., 2015; Brüggemann, 2021).

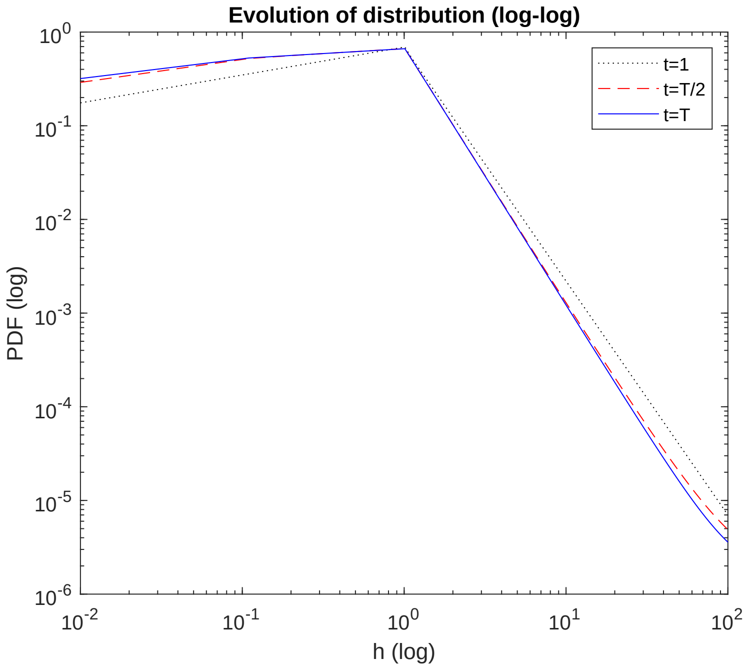


Figure 7: Change of the PDF after a shift in tax progressivity from $\tau_y = 18\%$ (baseline) to $\tau_y = 47\%$ (optimal in transition).

For an increase of taxes the drift of human capital m reduces which skews the distribution more to the left. The transitional behavior is the slowest in the top tails.

We account for such timing considerations by defining an adjusted welfare measure:

$$\begin{aligned} \hat{W}(\tau_y) = \int_0^T \exp(-\tilde{\rho}t) \int U(c,\ell) df_t(h) dt + \exp(-\tilde{\rho}T) \int U(c,\ell) df_T(h) \\ = \int_0^T \exp(-\tilde{\rho}t) W_t(\tau_y) dt + \exp(-\tilde{\rho}T) W_T(\tau_y). \end{aligned} \tag{68}$$

This welfare measure assumes a policy change in $t = 0$ and accounts for the change in the distribution discounted at some time preference rate $\tilde{\rho}$. In the numerical investigation we consider a time horizon $T = 70$ years for which the resulting distribution is virtually indistinguishable from the steady state distribution ($f_T(h) \approx f_{ss}(h)$). We then find the fixed level of tax progressivity that maximizes this welfare measure, upon a shift from the baseline degree of $\tau_y = 18\%$. This is found to be $\tau_y = 47\%$. This is higher than what is optimal when considering the steady state only ($\tau_y = 28\%$ as displayed in Table 3). Details on the numerical procedure are presented in Online Appendix G.

Figure 7 shows the distribution of human capital along the transition revealing its characteristic triangle pattern on a log-log scale. We have plotted the distribution at three select points in the transition, namely the one-year mark ($t = 1$), mid-transition ($t = T/2$) and end point ($t = T$). We consider the case of an increase in tax progressivity from the steady state level of $\tau_y = 18\%$ to the optimal level along the transition $\tau_y = 47\%$. As argued earlier, an increase in tax progressivity reduces human capital accumulation

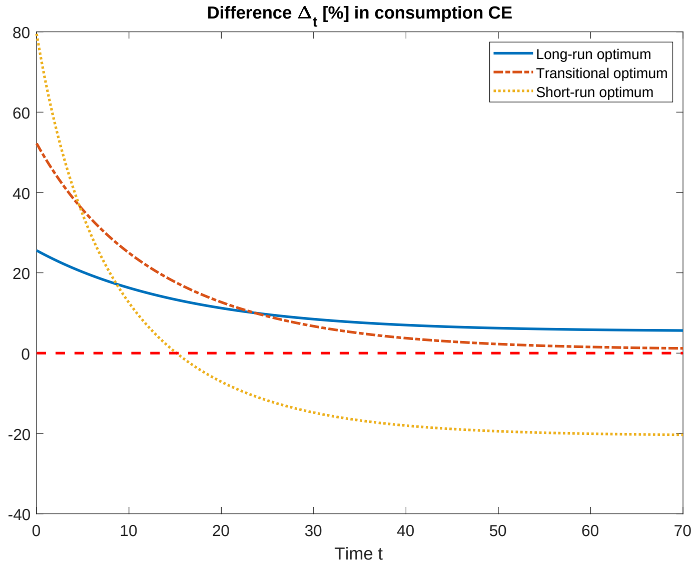


Figure 8: The time evolution of welfare in terms of change in CE consumption $\Delta_t(\tau_y)$ (relative to the baseline) for different optimal (utilitarian criterion) tax rates as reported in Table 11.

Short-run optimization aims at optimizing welfare at $t = 1$; whereas the long-run optimum optimization targets the terminal date ($t = T$). The transitional optimum optimizes discounting with a rate of $\rho = 5\%$.

m graphically spinning the distribution clockwise around the mode h_0 implying more bottom but less top inequality for human capital. As put forward in Gabaix et al. (2016) this behavior is the slowest in the distribution tails – especially at the very top.

In Figure 8 we show how the difference in certainty equivalent consumption $\Delta_t(\tau_y)$ (relative to the baseline with $\tau_y = 18\%$ under a once-for-all unexpected change in tax policy[33]) evolves in time for different tax rates optimal under different circumstances. The underlying optimal tax rates are displayed in Table 11. A shorter perspective promotes higher taxes since the welfare gains from redistribution are immediately reaped, while its adverse consequences – especially on aggregate human capital – only manifest themselves in the long run. More precisely, optimizing welfare only on impact ($t = 1$) would suggest tax progressivity as high as $\tau_y = 69\%$. While on impact such high progressivity produces very large gains, in the long run the situation is inferior to the baseline ($\Delta_T < 0$). This could be considered as the extreme case of a social planner that does not care about any future effects ($\tilde{\rho} \to \infty$). The other extreme covered throughout the paper can be considered as the case with $\tilde{\rho} = 0$ aimed at maximizing terminal

[33]We assume that the policy maker stays committed to a policy and do not allow for time-varying policies. As put forward in Heathcote et al. (2020a) a policy maker has an incentive to heavily redistribute earnings from already accumulated human capital in the short-run and to tax it less in the future to reduce adverse incentive effects. This insight holds also for taxation of physical capital where research suggests a decreasing time path for capital taxes (cf. e.g. Straub and Werning, 2020).

	$t=1$	$t=T$	Transition ($\tilde{\rho}=\rho=5\%$)
Utiltitarian ($\tilde{\gamma}=1$)	0.69	0.33	0.47
+ Government ($\phi>0$)	0.64	0.25	0.4
$\tilde{\gamma}=0$	0.52	0.00	0.00
$\tilde{\gamma}=2$	0.76	0.56	0.63
$\tilde{\gamma}\to\infty$	1	1	1

Table 11: Optimal tax progressivity for different time horizons
The shorter the perspective, the higher the optimal tax progressivity. Optimal tax progressivity is increasing with the curvature of the preferences of the social planner $\tilde{\gamma}$.

welfare.[34] A social planner with a positive finite discount rate – which we here equalize with the individual discount rate $\tilde{\rho}=\rho=5\%$ (cf. calibration in Table 2) – would choose an intermediate value of $\tau_y=47\%$. In the spirit of Bakis et al. (2015), one can consider the chosen discount rate as the degree of *government altruism*. The progressivity of tax rates considered to be optimal is thus an increasing function of the discount rate $\tilde{\rho}$.

All optimal progressivity levels are characterized by a higher value than prevailing in the benchmark (cf. Table 3). Moreover, they also vary by time path of welfare. In fact, optimal progressivity is decreasing in time suggesting that the planner frontloads welfare. In other words, a tax increase leads to an immediate downward shift in post-tax inequality which has positive welfare effects. In contrast, the pre-tax inequality is only slowly moving. The adverse consequences of this tax increase reducing human capital and hence the aggregate size of the pie available for redistribution mainly emerge in the long run. Hence, the path of welfare is declining in time. Following the same reasoning, a tax cut relative to the status quo would lead to an immediate negative shift, but an increasing welfare path in time.

As before, a preference for government consumption ($\phi>0$) lowers optimal tax progressivity, while it is increasing with the curvature of the preferences of the social planner $\tilde{\gamma}$ (cf. Table 11). Even considering the problem through the eyes of a representative agent ($\tilde{\gamma}=0$) would require a positive tax progressivity when only considering the shift upon impact (i.e., at $t=1$). In this case the optimal tax rate can be analytically characterized by $\tau_y=\frac{A}{A+\rho+p}>0$ with individuals enjoying the immediate increase in leisure. Under this tax regime individuals actually supply the same amount of labor they would supply in a laissez-faire economy with no taxation ($\tau_y=0$) and without on-the-job accumulation ($A=0$), i.e. $\ell=\chi^{-\frac{v}{1+v}}$. The high tax rates thus aligns the preferences of the households – who would like to accumulate human capital on-the-job – with the short run perspective of the social planner who does not account for long-run human capital accumulation.

[34]Note that the optimal tax reported here in the numerical exercise for $t=T$ is slightly higher than the one with the complete infinite horizon $T\to\infty$ presented in Section 3 as the distribution has not completely converged.

$\upsilon = 0.2$	$t = 1$	$t = T$	Transition ($\tilde{\rho} = \rho = 5\%$)
Utiltitarian ($\tilde{\gamma} = 1$)	0.74	0.45	0.56
+ Government ($\phi > 0$)	0.7	0.38	0.5
$\tilde{\gamma} = 0$	0.52	0.00	0.00
$\tilde{\gamma} = 2$	0.8	0.64	0.69
$\tilde{\gamma} \to \infty$	1	1	1
$\upsilon = 1$			
Utiltitarian ($\tilde{\gamma} = 1$)	0.65	0.22	0.38
+ Government ($\phi > 0$)	0.59	0.13	0.28
$\tilde{\gamma} = 0$	0.52	0.00	0.00
$\tilde{\gamma} = 2$	0.72	0.47	0.56
$\tilde{\gamma} \to \infty$	1	1	1

Table 12: Optimal tax progressivity for different time horizons with a variation of Frisch elasticity

In an economy that does not feature on-the-job learning ($A = 0$) optimal taxes from the perspective of the representative agent would still be zero (cf. Table 3). Therefore, the case with exogenous human capital calls for (slightly) lower tax progressivity than the optimization on impact ($t = 1$, cf. Table 11).

As before (cf. Online Appendix D) a higher Frisch elasticity – implying a more pronounced reaction of labor supply to tax changes – suggests a lower optimal tax progressivity (cf. Table 12). It is also interesting to note that for the extreme cases of a policy maker only concerned with the representative agent ($\tilde{\gamma} = 0$) or the least well-off ($\tilde{\gamma} \to \infty$) results remain unchanged for all $\upsilon > 0$ for a given time horizon.

F. Description of Numerical Algorithm in Steady State

1- Iterative Solution Method for the Hamilton-Jacobi-Bellman Equation

The HJB-equation is solved by a finite difference approach in accordance with Nuño and Moll (2018). The value function $V(a, h)$ is thus approximated on a finite grid. For human capital h, we employ a linear grid with constant step sizes Δh on $h \in \{h_1, \ldots, h_J\}$. To capture both the bottom end (featuring the largest chunk of the population) and the top end of the asset holdings $a \in \{a_1, \ldots, a_I\}$ a non-linear grid is deployed. Following Kindermann and Krueger (2022), we use $a_i = (a_I - a_1)\frac{(1+\zeta)^{i-1}-1}{(1+\zeta)^{I-1}-1} + a_1$ for which $\zeta > 0$ determines the convexity of the grid steering to what extent Δa_i is larger for higher values of a_i. For simplicity denote $V_{i,j} \equiv V(a_i, h_j)$ with $i \in \{1, \ldots, I\}$ and $j \in \{1, \ldots, J\}$. We approximate the derivative of optimal current value function V with respect to assets a by either forward or backward difference:

$$\frac{\partial}{\partial a}V_{i,j} \approx \partial_{a,F}V_{i,j} \equiv \frac{V_{i+1,j}-V_{i,j}}{\Delta a_i} \tag{69}$$

$$\frac{\partial}{\partial a}V_{i,j} \approx \partial_{a,B}V_{i,j} \equiv \frac{V_{i,j}-V_{i-1,j}}{\Delta a_i} \tag{70}$$

Whether the forward or backward difference is used depends on sign of savings $s_{i,j} = (wh_j\ell_{i,j})^{1-\tau_y}T + (1-\tau_K)ra_i - (1+\tau_C)c_{i,j}$ and is applied through a so-called upwind scheme. Similar to wealth the first-order derivative with respect to human capital h is also approximated by an upwind scheme. The second order derivative of optimal current value function V with respect to human capital h is approximated as follows:

$$\frac{\partial^2}{\partial h^2}V_{i,j} \approx \partial_{hh}V_{i,j} \equiv \frac{V_{i,j+1}+V_{i,j-1}-2V_{i,j}}{(\Delta h)^2}, \tag{71}$$

Recall the HJB-equation (28):

$$\rho V = U(c,\ell) + \frac{\partial V}{\partial h}\left(Aq(\ell)-\delta_h\right)h + \frac{\partial V}{\partial a}\left((wh\ell)^{1-\tau_y}T + (1-\tau_K)ra - (1+\tau_C)c\right) + \frac{1}{2}\sigma^2 h^2 \frac{\partial^2 V}{\partial h^2} + p(V(a,h_0)-V(a,h))$$

where

$$U(c,\ell) = \begin{cases} \frac{c^{1-\gamma}-1}{1-\gamma} - \chi\frac{\ell^{1+\frac{1}{\nu}}}{1+\frac{1}{\nu}}, & \text{if } \gamma \neq 1 \\ \ln c - \chi\frac{\ell^{1+\frac{1}{\nu}}}{1+\frac{1}{\nu}}, & \text{if } \gamma \to 1. \end{cases} \qquad q(\ell) = \begin{cases} \frac{\ell^{\theta}-1}{\theta}, & \text{if } \theta > 0, \\ \ln(\ell), & \text{if } \theta \to 0. \end{cases}$$

and hence $c = \left(\frac{\partial V}{\partial a}(1+\tau_C)\right)^{-\frac{1}{\gamma}}$ and

$$-\chi\ell^{\frac{1}{\nu}} + \frac{\partial V}{\partial h}Ahq'(\ell) + \frac{\partial V}{\partial a}(wh)^{1-\tau_y}(1-\tau_y)\ell^{-\tau_y}T = 0 \tag{72}$$

implicitly provides the labor supply decision. This non-linear root finding problem is solved by Newton's method with the initial guess $\ell^0 = \left(\frac{\frac{\partial V}{\partial h}Ah}{\chi}\right)^{\frac{\upsilon}{1+\upsilon(1-\theta)}}$ that would solve exactly for $\frac{\partial V}{\partial a} = 0$.

The upwind scheme approximating (28) is as follows

$$\frac{V_{i,j}^{t+1}-V_{i,j}^{t}}{\Delta} + \rho V_{i,j}^{t+1} = U(c_{i,j}^t,\ell_{i,j}^t) + \partial_{a,F}V_{i,j}^{t+1}s_{i,j,F}^t 1_{s_{i,j,F}^t>0} + \partial_{a,B}V_{i,j}^{t+1}s_{i,j,B}^t 1_{s_{i,j,B}^t<0} + \partial_h V_{i,j}^{t+1}(Aq(\ell_{i,j}^t)-\delta_h)h_j + \frac{\sigma^2 h_j^2}{2}\partial_{hh}V_{i,j}^{t+1} + p\left(V(a_i,h_0)^{t+1} - V_{i,j}^{t+1}\right) \tag{73}$$

where

$$s_{i,j,F}^t = (wh_j \ell_{i,j}^t)^{1-\tau_y} T + (1-\tau_K) r a_i - (\partial_{a,F} V_{i,j}^t (1+\tau_C))^{-1/\gamma},$$
$$s_{i,j,B}^t = (wh_j \ell_{i,j}^t)^{1-\tau_y} T + (1-\tau_K) r a_i - (\partial_{a,B} V_{i,j}^t (1+\tau_C))^{-1/\gamma}.$$

The numerical HJB in (73) can then be rewritten as

$$\frac{V_{i,j}^{t+1} - V_{i,j}^t}{\Delta} + \rho V_{i,j}^{t+1} = U(c_{i,j}^t, \ell_{i,j}^t) + V_{i-1,j}^{t+1} \varrho_{i,j} + V_{i,j}^{t+1} \beta_{i,j} + V_{i+1,j}^{t+1} \chi_{i,j} + V_{i,j-1}^{t+1} \xi_j + V_{i,j+1}^{t+1} \varsigma_{i,j} + p V^{t+1}(a_i, h_0) \tag{74}$$

where

$$\begin{aligned} &\left[\chi(\ell_{i,j}^t)^{1+\frac{1}{\nu}} - \partial_{a,F} V_{i,j}^t (wh)^{1-\tau_y} (1-\tau_y)(\ell_{i,j}^t)^{1-\tau_y} T\right] 1_{s_{i,j,F}^t > 0} \\ &\quad + \left[\chi(\ell_{i,j}^t)^{1+\frac{1}{\nu}} - \partial_{a,B} V_{i,j}^t (wh)^{1-\tau_y} (1-\tau_y)(\ell_{i,j}^t)^{1-\tau_y} T\right] 1_{s_{i,j,B}^t < 0} \\ &\quad + \left[\ell_{i,j}^t - \ell_{max}\right] 1_{s_{i,j,F}^t < 0, s_{i,j,B}^t > 0} = \partial_h V_{i,j}^t A h \left(1 - 1_{s_{i,j,F}^t < 0, s_{i,j,B}^t > 0}\right) q'(\ell_{i,j}^t) \ell_{i,j}^t \end{aligned} \tag{75}$$

$$\begin{aligned} c_{i,j}^t = &(\partial_{a,F} V_{i,j}^t (1+\tau_C) 1_{s_{i,j,F}^t > 0} + \partial_{a,B} V_{i,j}^t (1+\tau_C) 1_{s_{i,j,B}^t < 0} \\ &+ ((wh_j \ell_{i,j}^t)^{1-\tau_y} T + (1-\tau_K) r a_i)^{-\gamma} 1_{s_{i,j,F}^t < 0, s_{i,j,B}^t > 0})^{-\frac{1}{\gamma}} \end{aligned} \tag{76}$$

and

$$\begin{aligned} \varrho_{i,j} &= -\frac{s_{i,j,B}^t 1_{s_{i,j,B}^t < 0}}{\Delta a} \\ \beta_{i,j} &= -\frac{s_{i,j,F}^t 1_{s_{i,j,F}^t > 0}}{\Delta a} + \frac{s_{i,j,B}^t 1_{s_{i,j,B}^t < 0}}{\Delta a} - \frac{(Aq(\ell_{i,j}^t) - \delta_h) h_j}{\Delta h} - \frac{\sigma^2 h_j^2}{(\Delta h)^2} - p \\ \chi_{i,j} &= \frac{s_{i,j,F} 1_{s_{i,j,F}^t > 0}}{\Delta a} \\ \xi_j &= \frac{\sigma^2 h_j^2}{2(\Delta h)^2} \\ \varsigma_{i,j} &= \frac{(Aq(\ell_{i,j}^t) - \delta_h) h_j}{\Delta h} + \frac{\sigma^2 h_j^2}{2(\Delta h)^2}. \end{aligned}$$

The state lower boundary constraint $a \geq a_{min}$ is satisfied by letting $s_{1,j,B} = 0$. Equivalently, the upper-boundary constraint on assets $a \leq a_{max}$ is implemented by setting $s_{I,j,F} = 0$. Subsequently, the boundary current-value functions $V_{0,j}^{t+1}$ and $V_{I+1,j}^{t+1}$ are never used. Consequently, the boundary conditions on human capital h are the following:

$$\frac{\partial V(a, h_{min})}{\partial h} = \frac{\partial V(a, h_{max})}{\partial h} = 0.$$

Hence, the boundary equations of (74) along the J-dimension become:

$$\frac{V_{i,1}^{t+1}-V_{i,1}^{t}}{\Delta}+\rho V_{i,1}^{t+1}=U(c_{i,1}^{t},\ell_{i,1}^{t})+V_{i-1,1}^{t+1}\varrho_{i,1}+V_{i,1}^{t+1}(\beta_{i,1}+\xi_1)+V_{i+1,1}^{t+1}\chi_{i,1}+V_{i,2}^{t+1}\varsigma_{i,1}+pV^{t+1}(a_i,h_0),$$
$$\frac{V_{i,J}^{t+1}-V_{i,J}^{t}}{\Delta}+\rho V_{i,J}^{t+1}=U(c_{i,J}^{t},\ell_{i,J}^{t})+V_{i-1,J}^{t+1}\varrho_{i,J}+V_{i,J}^{t+1}(\beta_{i,J}+\varsigma_{i,J})+V_{i+1,J}^{t+1}\chi_{i,J}+V_{i,J-1}^{t+1}\xi_J+pV^{t+1}(a_i,h_0).$$

We can thereby reformulate the $I\times J$ system of equations in (74) in matrix form as

$$\frac{\mathbf{V}_{t+1}-\mathbf{V}_t}{\Delta}+\rho\mathbf{V}_{t+1}=\mathbf{U}_t+(\mathbf{A}_t+\mathbf{P})\mathbf{V}_{t+1}$$

where,

.

$$\mathbf{A}_t=\begin{bmatrix}
\beta_{1,1}+\xi_1 & \chi_{1,1} & 0 & \cdots & 0 & \varsigma_{1,1} & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0\\
\varrho_{2,1} & \beta_{2,1}+\xi_1 & \chi_{2,1} & 0 & \cdots & 0 & \varsigma_{2,1} & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0\\
0 & \varrho_{3,1} & \beta_{3,1}+\xi_1 & \chi_{3,1} & 0 & \cdots & 0 & \varsigma_{3,1} & \cdots & 0 & 0 & 0 & 0 & 0 & 0\\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & 0\\
0 & \cdots & 0 & \varrho_{I,1} & \beta_{I,1}+\xi_1 & \chi_{I,1} & 0 & \cdots & 0 & \varsigma_{I,1} & 0 & \cdots & 0 & & \\
\xi_2 & 0 & \cdots & 0 & \varrho_{1,2} & \beta_{1,2} & \chi_{1,2} & 0 & \cdots & 0 & \varsigma_{1,2} & 0 & \cdots & 0 & \\
0 & \xi_2 & 0 & \cdots & 0 & \varrho_{2,2} & \beta_{2,2} & \chi_{2,2} & 0 & \cdots & 0 & \varsigma_{2,2} & 0 & \cdots & 0\\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0\\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \varsigma_{I,J-1}\\
0 & \cdots & \cdots & \cdots & 0 & \xi_J & 0 & \cdots & 0 & \varrho_{1,J} & \beta_{1,J}+\varsigma_{1,J} & \chi_{1,J} & 0 & \cdots & 0\\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0\\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0\\
0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \xi_J & 0 & \cdots & 0 & \varrho_{I-1,J} & \beta_{I-1,J}+\varsigma_{I-1,J} & \chi_{I-1,J}\\
0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \xi_J & 0 & \cdots & 0 & \varrho_{I,J} & \beta_{I,J}+\varsigma_{I,J}+\chi_{I,J}
\end{bmatrix}_{IJ\times IJ}$$

$$\mathbf{P}=\begin{bmatrix} \underbrace{0}_{I\times j'(h_0\text{'s index})} & \begin{vmatrix} P_{I\times J}\\ P_{I\times J}\\ \vdots\\ P_{I\times J}\end{vmatrix} & 0 \end{bmatrix}_{IJ\times IJ} \text{ with } P_{I\times J}=p\begin{bmatrix} 1 & 0 & 0\cdots 0\\ & \ddots & \vdots \quad \vdots\\ 0 & 1 & 0\cdots 0\end{bmatrix} \text{ if } J>I \text{ or } P_{I\times J}=p\begin{bmatrix} 1 & 0\\ & \ddots\\ 0 & 1\\ 0 & \cdots 0\\ \vdots & \vdots\\ 0 & \cdots 0\end{bmatrix} \text{ if } I>J,$$

$$\mathbf{V}_t=\begin{bmatrix} V_{1,1}^t\\ V_{2,1}^t\\ \vdots\\ V_{1,2}^t\\ V_{2,2}^t\\ \vdots\\ V_{I-1,J}^t\\ V_{I,J}^t\end{bmatrix},\mathbf{U}_t=\begin{bmatrix} U(c_{1,1}^t,\ell_{1,1}^t)\\ U(c_{2,1}^t,\ell_{2,1}^t)\\ \vdots\\ U(c_{1,2}^t,\ell_{1,2}^t)\\ U(c_{2,2}^t,\ell_{2,2}^t)\\ \vdots\\ U(c_{I-1,J}^t,\ell_{I-1,J}^t)\\ U(c_{I,J}^t,\ell_{I,J}^t)\end{bmatrix}.$$

The parameter t can considered a *virtual* time step, for which the criterion $||\mathbf{V}_{t+1} - \mathbf{V}_t|| < \epsilon$ is employed to measure sufficient convergence to the infinite horizon solution.

The toolbox suggested in Achdou et al. (2022) and Nuño and Moll (2018) works well for quasi-linear problems, however, becomes problematic in non-linear setups as the one discussed here. We therefore follow Gomez (2015) who proposes the use of the Pseudo-Transient continuation method (Coffey et al., 2003) in an economic setup. It requires another inner loop ($n = 1, \cdots, N$) to solve the HJB in which the meta-parameter Δ (the *time step*) is dynamically adjusted according to:

$$\Delta^n = \Delta^{n-1}||F(\mathbf{V}_{t+1}^{n-1})||/||F(\mathbf{V}_{t+1}^{n})||, \tag{77}$$

for which $F(\mathbf{V}) = (\mathbf{A}+\mathbf{P}-\rho\mathbf{I})\mathbf{V}+\mathbf{U}$ is the error in the HJB equation which we evaluate in an infinity norm. It implies that for low HJB errors the value of Δ is increased (and vice versa) leading to a more stable convergence pattern.

The updating in the inner loop is given by:

$$\mathbf{V}_{t+1}^{n+1} = \left(\left(\frac{1}{\Delta^n} + \rho\right)\mathbf{I} - \mathbf{A}_t\right)^{-1}\left(\mathbf{U}_t + \frac{\mathbf{V_{t+1}^{n}}}{\Delta^n}\right) \tag{78}$$

It follows from applying Newton's method on the HJB. In case the number of inner loops are restricted to $N = 1$, the algorithm of Achdou et al. (2022) with a fixed Δ is nested. We set $N = 5$.

Hence, the HJB-equation in (28) is solved as follows:

1. Guess an initial guess for the current-value function $V_{i,j}^0$. Set $n = t = 0$. We use $V_{i,j}^0 = \frac{(h_j+a_i)^{1-\gamma}}{1-\gamma} + K_2$ for which the latter is defined in Appendix A.
2. Compute $\partial_h V_{i,j}^t$, $\partial_{hh} V_{i,j}^t$, $\partial_{a,B} V_{i,j}^t$ and $\partial_{a,F} V_{i,j}^t$ using (69) to (71).
3. Find $\ell_{i,j}^t$ using (75).
4. Calculate $c_{i,j}^t$ using (76).
5. Iterate (at most) N times (or until convergence) on 78 to get $\mathbf{V}_{t+1}^{N+1}$ also updating Δ_n according to 77.
6. If $\mathbf{V}_{t+1}^{N+1}$ is not significantly different from $\mathbf{V}_t^{N+1}$, (here the infinity norm guarantees global convergence), then stop. Otherwise, update $\mathbf{A}_{t+1}$ and $\mathbf{U}_{t+1}$ and repeat from step 2.

2- Solving the Kolmogorov-Forward Equation

The converged iterative HJB-equation, and the coefficient matrix $\mathbf{A} = \lim_{t\to\infty} \mathbf{A}_t$ that is yielded can be used in solving the KF-equation. In practice, $\mathbf{A}$ will be the result of the last iteration of $\mathbf{A}_t$ in the HJB-solving iterative scheme outlined above. Thus, we employ yet again an upwind finite difference scheme. The KF-equation is given by

$$
\begin{aligned}
-\frac{\partial}{\partial a}\left[s(a,h)f(a,h)\right] - \frac{\partial}{\partial h}\left[\left(Aq(\ell)-\delta_h\right)hf(a,h)\right] \\
+\frac{1}{2}\frac{\partial^2}{\partial h^2}\Big(\sigma^2 h^2 f(a,h)\Big) + p(f(a,h_0)-f(a,h)) = 0 \\
\text{such that } \int\int f(a,h)dhda = 1 \text{ and } s(a,h) \text{ as in (28).}
\end{aligned}
$$

Using the the notation $f_{i,j} = f(a_i, h_j)$, the system of equations becomes

$$
\begin{aligned}
&-\frac{f_{i,j}s^t_{i,j,F}1_{s^t_{i,j,F}>0} - f_{i-1,j}s^t_{i-1,j,F}1_{s^t_{i-1,j,F}>0}}{\Delta a} - \frac{f_{i+1,j}s^t_{i+1,j,B}1_{s^t_{i+1,j,B}<0} - f_{i,j}s^t_{i,j,B}1_{s^t_{i,j,B}<0}}{\Delta a} \\
&\quad -\frac{f_{i,j}\Big(Aq(\ell^t_{i,j})-\delta_h\Big)h_j - f_{i,j-1}\Big(Aq(\ell^t_{i,j-1})-\delta_h\Big)h_{j-1}}{\Delta h} + \frac{\sigma^2(f_{i,j+1}+f_{i,j-1}-2f_{i,j})}{2(\Delta h)^2} \\
&\qquad + p(f(a_i,h_0)-f_{i,j}) = 0
\end{aligned}
$$

or put in terms of already-defined coefficients

$$
\beta_{i,j}f_{i,j} + \chi_{i-1,j}f_{i-1,j} + \varrho_{i+1,j}f_{i+1,j} + \xi_j f_{i,j+1} + \varsigma_{i,j}f_{i,j-1} = -pf(a_i,h_0).
$$

This is again an $I \times J$ system of linear equations given in matrix form by

$$
(\mathbf{A}^T + \mathbf{P})\mathbf{f} = \mathbf{b},
$$

(with T being the transpose operator) where,

$$\mathbf{f} = \begin{bmatrix} f_{1,1} \\ f_{2,1} \\ \vdots \\ f_{1,2} \\ f_{2,2} \\ \vdots \\ f_{I-1,J} \\ f_{I,J} \end{bmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

This step is then concluded with normalization:

$$f_{i,j} := \frac{f_{i,j}}{\sum_{i=1}^{I} \sum_{j=1}^{J} f_{i,j} \Delta a \Delta h}.$$

3- Finding Equilibrium Values

We employ a relaxation method to find the aggregate levels of physical capital (K), human-capital-augmented labor supply (HL), labor income taxation T, and consequently the equilibrium wage w and real rates of return r. These equilibrium values in a perfectly competitive economy are given by setting demand equal to supply

$$r = \alpha \frac{Y}{K} - \delta_k \tag{79}$$

$$w = (1-\alpha) \frac{Y}{HL} \tag{80}$$

where $Y = \Xi K^{\alpha} (HL)^{1-\alpha}$ is total output.

We first choose a relaxation parameter $\epsilon \in (0,1)$. Second we guess an initial value for physical capital K^0, human-capital-augmented labor supply $(HL)^0$, labor income taxation T^0 and set $t = 0$. Then

1. Derive r_t and w_t as function of K_t, $(HL)_t$ and T_t in accordance with (79) and (80).
2. Use r_t and w_t in order to solve the HJB-equation in Step 1 to retrieve estimates of the value function V_t, consumption c_t and labor supply ℓ_t.

3. Granted c_t and ℓ_t, solve the KF-equation in Step 2 and compute the density distribution g^n.

4. Evaluate the updated aggregate values through

$$
\begin{aligned}
(SK)_t &= \sum_{i=1}^{I}\sum_{j=1}^{J} a_i f_{i,j}\Delta a\Delta h \\
(SHL)_t &= \sum_{i=1}^{I}\sum_{j=1}^{J} h_j \ell_{i,j} f_{i,j}\Delta a\Delta h \\
(ST)_t &= w^{\tau_y}\frac{(SHL)_t}{\sum_{i=1}^{I}\sum_{j=1}^{J}(h_j\ell_{i,j})^{1-\tau_y} f_{i,j}\Delta a\Delta h}
\end{aligned}
$$

5. Relaxedly $(0 < \omega < 1)$ update the the aggregate values through

$$
\begin{aligned}
K_{t+1} &= \omega(SK)_t + (1-\omega)K_t \\
(HL)_{t+1} &= \omega(SHL)_t + (1-\omega)(HL)_t \\
T_{t+1} &= \omega(ST)_t + (1-\omega)T_t.
\end{aligned}
$$

If the update in K_{t+1}, $(HL)_{t+1}$ and T_{t+1} is small compared to previous iteration, stop. Otherwise, set t to $t+1$ and repeat from step 1.

G. Computational algorithm in the transition

1. Compute the steady state $\mathbf{f}_0$ under the prevailing policy τ_y.

2. Compute the time-invariant transition matrix $\mathbf{A}$ under a new tax policy τ_y'.

3. *Forward simulation:* Simulate the distribution forward from $t = 0$.

   For some transition matrices $\mathbf{A}^T + \mathbf{P}$ the distribution evolves in time according to:

$$
\frac{d\mathbf{f}}{dt} = (\mathbf{A}^T + \mathbf{P})\mathbf{f}. \tag{81}
$$

   This can be approximated by a finite difference (Δ) scheme and a backward Euler equation (implicit method):

$$
\frac{\mathbf{f}_{t+1} - \mathbf{f}_t}{\Delta} = (\mathbf{A}^T + \mathbf{P})\mathbf{f}_{t+1}. \tag{82}
$$

We can reformulate this equation as follows:

$$\mathbf{f}_{t+1} = \left[\mathbf{I} - (\mathbf{A}^T + \mathbf{P})\Delta\right]^{-1} \mathbf{f}_t. \tag{83}$$

Starting with the initial distribution $\mathbf{f}_0$ we compute the subsequent distributions using a finite difference scheme running forward in time. We consider a discrete time step of $\Delta = 0.1$.

H. Additional graphs and analysis of the GE model

Figure 9 illustrates the decision rules and utility as functions of human capital h and assets a in the stationary equilibrium for the baseline calibration ($\tau_y = 0.18$). Labor supply is increasing in human capital, while decreasing in assets. Consumption is strictly increasing in both states being more sensitive to asset holdings. The joint distribution reveals the characteristic triangle patter in the human capital dimension.

Additionally, we consider a model absent human capital accumulation. In order to match the moments the model absent human capital accumulation is recalibrated. Table 13 shows that the recalibrated version matches the targeted moments as well as the original calibration reported in Table 5. To abstract from human capital accumulation, we set $A = 0$ and $m = -\delta_h = 2.8\%$ as suggested in Table 1. Due to relative risk-aversion being larger than one ($\gamma > 1$), there is a dispersion of hours worked making the human capital distribution not identical to the income distribution. This dispersion is, however, minuscule allowing for the variance of shocks to human capital σ^2 being very close to the PE model (cf. Table 1) and substantially higher than in the GE model. Other than that, the calibrated values are very close to the ones from the baseline. All other values are kept as reported in Table 4.

Category	Name	Value	Goal	Data	Model
Preferences	Disutility labor χ	0.19	Average labor supply	1.4	1.40
Income process	Volatility σ^2	7.8%	Gini income	0.58	0.58
Production	Total Factor Productivity Ξ	0.28	Wealth / output ratio	3	3.01
	Depreciation δ_k	7.15%	Interest rate	4.5%	4.8%

Table 13: Calibrated parameters with targeted moment and fit for the model with exogenous human capital.

The calibrated model fits the targeted moments very well.

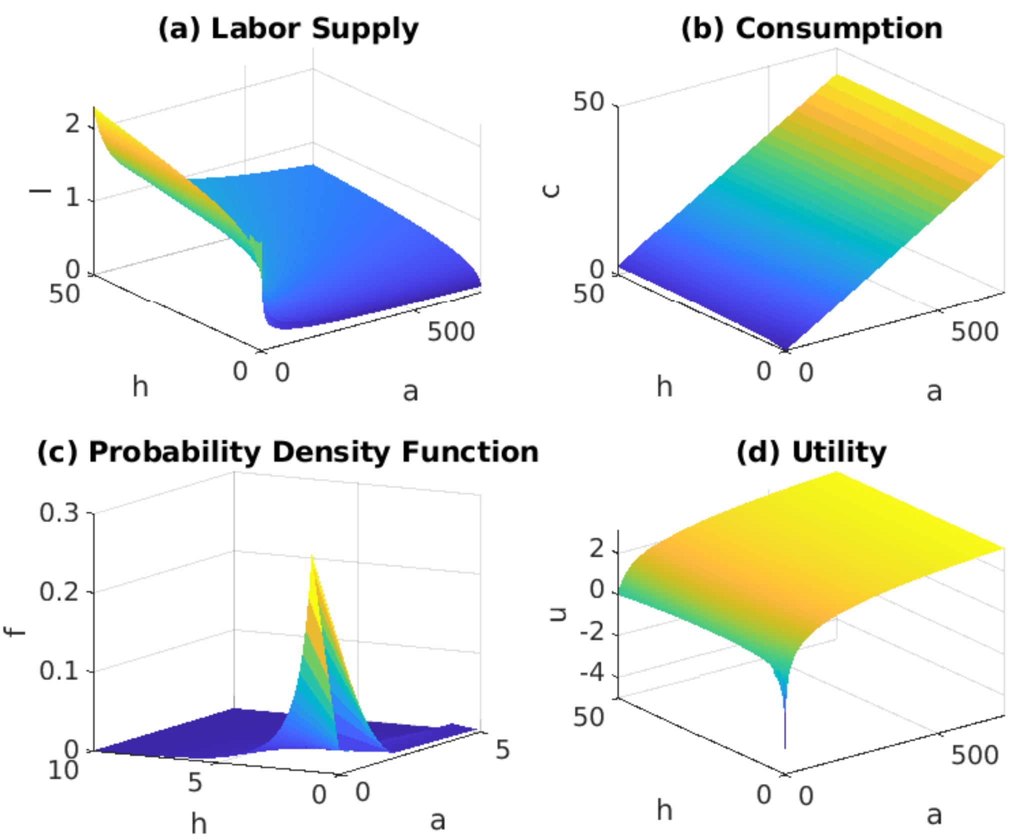


Figure 9: Decision Rules and Utility in the Baseline General Equilibrium Model.

Panel (a) shows the labor supply decisions ℓ, panel (b) the consumption c, panel (c) the probability density function f and panel (d) the utility U as functions of human capital h and assets a in the stationary equilibrium for the baseline calibration ($\tau_y = 0.18$).