# Optimal Progressive Income Taxation for Skewed Business Cycles

*By* William R. Bennett*

May 20, 2025

*I provide new estimates of forecaster inattention and information rigidity related to the sticky information model of expectation formation. While most papers use aggregate-level regressions or model calibrations to estimate the amount of information rigidity, I employ a more granular estimation strategy using micro-level data from the US Survey of Professional Forecasters (SPF). I provide evidence that the true amount of forecaster inattention is much smaller than previously thought. I also document novel state-dependence and time series facts. My results imply that some other, stronger source of information rigidity must exist to account for the discrepancy between the aggregate- and micro-level results. A model accounting for information rigidity should not use sticky information as neither its only nor main mechanism, as doing so could result in incorrect model predictions.*

## I. Introduction

For many countries, one of the government's responsibilities is to promote the general welfare of its populace. A common way a government does this is by engaging in income redistribution financed by progressive income taxes. An extensive literature has studied how redistribution and income taxes should be structured, as while taxation can both increase redistribution and help insure against idiosyncratic income shocks, it can also create distortions by disincentivizing some individuals from working and saving. The literature that accounts for the heterogeneity of households usually excludes aggregate fluctuations, meaning their calculated optimal level of progressivity does not factor in business cycles. However, it is well documented that the distribution of idiosyncratic income shocks differs between economic expansions and recessions.[1] Because income growth risk—and to an extent inequality—vary over the business cycle, there are potential welfare gains to having the progressivity of the government's tax and transfer system vary as well.

To investigate this possibility I conduct a quantitative analysis of optimal fiscal policy in a

*Department of Economics, University of Houston (email: wrbennett@uh.edu)

[1]See Storesletten, Telmer, and Yaron (2004) and Guvenen, Ozkan, and Song (2014) for classic examples.

heterogeneous-agent incomplete markets (HAIM) economy with aggregate shocks. I expand upon the standard model in two main ways.

First, the government has access to a log-linear tax and transfer schedule for income, as opposed to the usually imposed flat tax rate and lump sum transfer. This tax and transfer schedule will be represented by a simple two parameter function of income where the government's choice variable will be the parameter controlling the progressivity of the schedule. Moreover, the government can have different schedules (functions) for different aggregate states of the economy. Specifically, I model business cycles as a two-state Markov process representing expansion and recession and the government can have separate tax and transfer schedules for each state.[2]

Second, the standard AR(1) individual productivity process is augmented to incorporate recent findings in the income growth literature. As documented in multiple works by Fatih Guvenen, such as Guvenen, Ozkan, and Song (2014) and Guvenen et al. (2021), analysis of US social security administration (SSA) data reveals income growth has substantial skewness and kurtosis (fat-tails). High kurtosis means that experiencing a very large increase or decrease in income is much more likely then previously thought. High skewness means that while the majority of individuals will experience positive income growth and be on the right side of the distribution, the overall shape of the distribution is stretched to the left so that the unlucky individuals who do happen to experience a negative shock will get a large decrease relative to the increases of the people on the right side.[3] Furthermore, the amount of skewness is on average twice as high in a recession than in an expansion.[4] This is in contrast to the standard model, which assumes income growth has zero skewness and zero kurtosis, with innovations coming from a normal distribution. The standard model also doesn't incorporate any state-dependence on the distribution of the innovations.

Figure 1 visually summarizes this phenomenon. While productivity growth is approximately normal when switching from recession to expansion (the RE transition), the other three possible aggregate transitions all show significant left-skewness and fat-tails. While the left-tails of the expansion to expansion (EE) and recession to recession (RR) productivity growth distributions are visible, it is impossible to adequately overlay the left-tail of the expansion to recession (ER) distribution on the same graph. While it appears that the ER distribution is just heavily concentrated around zero, it in fact has the highest amount of skewness and kurtosis, with a significant mass below -2 where the other distributions don't. For exposition, a draw of -2 from one of these distri-

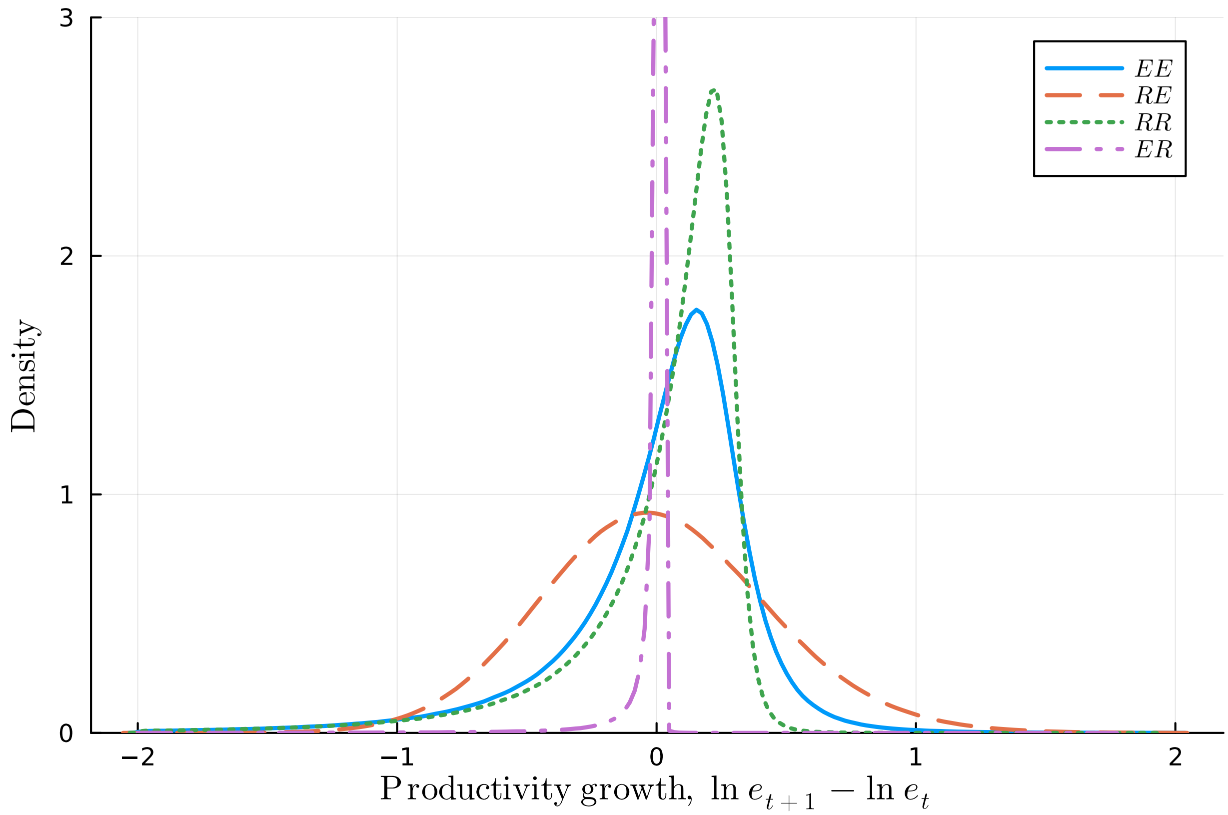
[2]Modeling aggregate fluctuations as a discrete Markov process has a detailed history. See Hamilton (1990) for a classic application of regime switching estimation. Krusell and Smith (1998) uses a two state process for both aggregate and individual productivities due to computational restrictions, but I am able to use a finer grid for individual productivity and still do optimal policy analysis.

[3]Holding the mean at zero and variance constant.

[4]The mean is also different between the two states, while variance and kurtosis are about the same.

butions means that annual income changes by $e^{-2} - 1$, i.e. it falls by 86%. This means that only when entering a recession does a significant proportion of households experience 90%-100% declines in labor earnings. The start of a recession is also marked by a drastic decline in any meaningful productivity increases.

Figure 1—State-Dependent Skewed and Fat-Tailed Productivity Growth Innovations




**Notes:** This figure plots the four aggregate state-dependent distributions for annual log productivity change. They are normalized to each have a mean of zero to focus on the distributions' shapes. They are calibrated in section III.A to match the autocorrelation, mean, standard deviation, standardized skewness, and standardized excess kurtosis of income growth in expansions and recessions. In the legend, the first letter for a line is the current aggregate state, and the second letter is the aggregate state next period. While truncated for the graph, the densities of the tails for most of the distributions stay significantly above zero, especially *ER*'s left-tail. *ER*'s density around zero is also cut-off at the top but extends significantly higher than three.

The skewness and fat-tails of productivity growth shocks have *a priori* unclear implications for optimal tax progressivity. *Ceteris paribus* and holding the mean and variance constant, excess kurtosis likely exacerbates income/consumption inequality, as the poor become poorer and the rich become richer despite there being a larger middle class. This would lead to greater desired redistribution and tax progressivity. However, the dominant left tail caused by negative skewness means there are many more severely poor individuals than there are richer ones, so much of the redistribution won't come from the super-rich, but instead from the larger middle class.

Using my model, I quantitatively solve for the socially optimal level of income tax progressivity and analyze the effects of state-dependent income growth skewness and progressivity. I calibrate the model using US data and study these effects on the aggregate variables as well as the responses of the heterogeneous agents via a simulated sequence of expansions and recessions.

Models with heterogeneous agents and aggregate shocks are typically intractable or very computationally intensive to solve due to the distribution of agents being an infinite-dimensional object that must be tracked over time. This makes doing quantitative policy analysis often extremely difficult or impossible.[5] To circumnavigate this issue, I utilize recent advancements in computational economics and adapt them to regime-switching processes to more quickly solve my model, allowing for optimal policy analysis.[6]

I find that the optimal tax and transfer policy is to have moderately more progressivity in expansions than in recessions, yet both should be the lower than the current (and historically average) US level. Furthermore, when solving the model with skewed business cycles, I find optimal progressivity is significantly positive, but if income growth shocks come from normal distributions than the income tax should be approximately flat/proportional.

Several papers in the taxation literature, such as Conesa and Krueger (2006) and Conesa, Kitao, and Krueger (2009), and Heathcote, Storesletten, and Violante (2020) look at optimal income tax progressivity, but do so by only analyzing policy in the steady state or by shutting down aggregate shocks. Others look at one time policy changes and transitional effects, such as in Krueger and Ludwig (2013), Bakis, Kaymak, and Poschke (2015), Ferrière et al. (2021), Dyrda and Pedroni (2023), and Boar and Midrigan (2021). A related literature looks at optimal progressive taxation in models with aggregate shocks, but restricts the tax function to a (usually fixed) flat tax rate with a lump sum transfer, such as Bhandari et al. (2021) and Angelopoulos, Asimakopoulos, and Malley (2019).

Closer to this paper, McKay and Reis (2021) solves for the optimal non-varying income tax progressivity in a model with aggregate fluctuations and unemployment and Zoi (2020) solves for the optimal path of progressivity given a one time shock to aggregate productivity with nonconstant idiosyncratic variance. Unlike these (and the previously mentioned) papers, I model aggregate productivity as a two state process to more accurately account for the patterns of fat-tailed cyclical skewness of income risk observed in the SSA data, as well as take into consideration the government's possible inabilities to frequently change the tax system.

[5]See Krusell and Smith (1998) for a classic discussion.
[6]Specifically the techniques of Boppart, Krusell, and Mitman (2018) and Auclert et al. (2021).

The rest of this paper is organized as follows. Section II presents the model, its agents, and its equilibrium. Section III explores quantitative exercises including calibration, computation, and results. Section IV concludes. V contains appendices.

## II. The Model

The quantitative model starts out similarly to the heterogeneous-household models of Aiyagari (1994) and Krusell and Smith (1998). Households face uninsurable productivity risk due to a no-borrowing constraint under incomplete markets. I depart from this benchmark in two important ways. First, I model aggregate uncertainty in the form of a two-state regime switching process that affects the distribution of productivity shocks. Second, the government has a log-linear tax function that it uses to finance required government expenditures and facilitate redistribution.

### A. *Households*

The economy is populated by a continuum of households of measure 1. Facing both idiosyncratic productivity risk and aggregate uncertainty, households make consumption, savings, and labor decisions across time. Households are indexed by $i$ in the unit interval. The idiosyncratic state of a household consists of its productivity $e_{it}$ and capital savings $k_{it}$. Households derive utility from consumption, $c_{it}$, and disutility from working, $h_{it}$, according to their lifetime expected utility function,

$$U_i = \mathbf{E}_0 \sum_{t=1}^{\infty} \beta^t u(c_{it}, h_{it}), \tag{1}$$

and their contemporaneous utility function,

$$u(c_{it}, h_{it}) = \frac{\left(c_{it} - \psi \frac{h_{it}^{1+\varphi}}{1+\varphi}\right)^{1-\theta}}{1-\theta}. \tag{2}$$

where $\beta$ is the discount factor, $\theta$ controls the coefficient of relative risk aversion and the inverse of the elasticity of intertemporal substitution, $\varphi$ is related to the inverse Frisch elasticity of labor supply, and $\psi$ is a scaling parameter capturing the relative importance of leisure to consumption.

Households are subject to a budget constraint,

$$c_{it} + k_{it+1} = \lambda_t (W_t e_{it} h_{it})^{1-\tau_t} + (1 + R_t - \delta) k_{it}, \tag{3}$$

and a borrowing constraint,

$$k_{it+1} \geq \underline{k}, \tag{4}$$

where $W_t$ is the wage rate of labor, $R_t$ is the rental rate of capital, $\delta$ is the depreciation rate of capital, $\underline{k}$ is the borrowing constraint, and $\lambda_t$ and $\tau_t$ are fiscal policy variables set by the government that control the level and progressivity of the income tax, respectively.

Let $y_{it}^0 \equiv w_t e_{it} h_{it}$ be a household's gross labor income ("pre-tax income"). Then a household's net labor income after taxes and transfers occur ("post-tax income") is given by $y_{it} \equiv f(y_{it}^0) = \lambda_t (w_t e_{it} h_{it})^{1-\tau_t}$, where $f$ is the log-linear tax function. This functional form for net income is a common representation in the literature of a progressive income tax system.[7] I will briefly describe some of the function's properties. $\tau$ is the "progressivity parameter" while $\lambda$ is the "level parameter". If $\tau = 0$, the the tax system is neither progressive nor regressive and each household faces the same average and marginal tax rates of $1 - \lambda$. Any $\tau$ between (0,1] makes the tax and transfer system progressive in the sense that the marginal tax rate is larger than the average tax rate for any income level. Depending on $\lambda_t$, $\tau_t$, and $y_{it}^0$ the household may have more, less, or the same amount of income before and after taxes and transfers. Specifically, there is a threshold $\bar{y}_t \equiv \lambda_t^{\frac{1}{\tau_t}}$ such that if $y_{it}^0 < \bar{y}_t$, then $y_{it}^0 < y_{it}$, and vis versa. If $\tau = 1$, then the system is perfectly progressive, as every household will have the same post-tax income regardless of their pre-tax income. Finally, if $\tau < 0$, then the tax system is regressive, as marginal rates decline with income.

### B. *Aggregate Fluctuations*

The aggregate state $z$ follows an exogenous Markov process with two possible states and four transition probabilities:

$$\Pi_{zz'} = \Pr(z_{t+1} = z' | z_t = z). \tag{5}$$

The two states are $z = E$, representing the economy being in an expansion, and $z = R$, representing the economy being in a recession. In this model there is not a notion of total factor productivity (TFP). Instead, aggregate shocks only indirectly affect the economy through their impact on 1) the distribution of shocks to individual productivity, and 2) potential state-dependent government tax policy.

[7] See Feldstein (1973) for a classic example and Heathcote, Storesletten, and Violante (2017) for a modern one.

### C. Individual Productivity

A large majority of macroeconomic models similar to this one have productivity shocks come from normal distributions. Unfortunately, doing this imposes a restriction of zero skewness and kurtosis on productivity/income growth, which contradicts the recent literature results on US social security administration income data. To allow income growth to be skewed and fat-tailed, the normal distribution restriction must be lifted.

In empirical work such as Guvenen, Ozkan, and Song (2014), Guvenen et al. (2021), and similar papers, each assume aggregate states have income growth shocks come from a mixture of two (sometimes three) normal distributions. The main appeal of these mixture distributions is its potential interpretation; with some probability, you keep your current job and experience a small change to income (such as a wage raise), represented by a small variance for that specific normal distribution. With one minus that probability you could lose or quit your job and either fail to quickly find another or succeed in finding a much better one, represented by a large variance for that specific normal distribution. These probabilities being dependent on the aggregate state allows the likelihood of drawing from the higher variance normal distribution to be higher in recessions.[8]

While having a nice interpretation, mixture models have a drawback when incorporated in computational models. Assuming a mixture of only two normal distributions, the first (relatively minor) issue is that it has five parameters (for each of the four aggregate transition possibilities): the probability of which normal to draw from, two mean terms, and two variance terms. This is one more than needed to control for four moments (mean, variance, skewness, and kurtosis), and can worsen the performance of moment matching algorithms when discretizing the productivity process. The second (relatively major) issue is if the true distribution posses considerably fat-tails, a mixture of two normal distributions will have difficulty capturing those tails. It would require a mixture of three or more to get a good approximation, in which case the number of required parameters becomes restrictively high and interpretation is lost.

To avoid these issues, I have the state-dependent distribution come from a four-parameter function that can more easily create skewness and fat-tails. I assume the log of individual productivity $e_{it}$ follows an AR(1) process with the distribution of innovations dependent on the transitioning of the

[8]The means of the normal distributions are also different but their discussion is removed for brevity.

aggregate state:

$$\ln(e_{it}) = \rho \ln(e_{it-1}) + \nu_{it}, \tag{6}$$

$$\nu_{it} \sim \begin{cases} S_U(\xi_{EE}, \zeta_{EE}, \gamma_{EE}, \eta_{EE}) & \text{if } z_t = z_E \text{ and } z_{t-1} = z_E \\ S_U(\xi_{RE}, \zeta_{RE}, \gamma_{RE}, \eta_{RE}) & \text{if } z_t = z_E \text{ and } z_{t-1} = z_R \\ S_U(\xi_{ER}, \zeta_{ER}, \gamma_{ER}, \eta_{ER}) & \text{if } z_t = z_R \text{ and } z_{t-1} = z_E \\ S_U(\xi_{RR}, \zeta_{RR}, \gamma_{RR}, \eta_{RR}) & \text{if } z_t = z_R \text{ and } z_{t-1} = z_R \end{cases}, \tag{7}$$

where $\rho$ is the persistence/autocorrelation of log productivity and $\nu_{it}$ is its innovation. $S_U$ is the Johnson's $S_U$-distribution, and the multiple $\xi$, $\zeta$, $\gamma$, and $\eta$ control the mean, variance, skewness, and kurtosis of the distributions, respectively. See Appendix V.A for more details about the Johnson's $S_U$-distribution.

The counterfactual scenario analyzed later will be to restrict the innovations to come from normal distributions:

$$\nu_{it} \sim \begin{cases} \mathcal{N}(\mu_{EE}, \sigma^2_{EE}) & \text{if } z_t = z_E \text{ and } z_{t-1} = z_E \\ \mathcal{N}(\mu_{RE}, \sigma^2_{RE}) & \text{if } z_t = z_E \text{ and } z_{t-1} = z_R \\ \mathcal{N}(\mu_{ER}, \sigma^2_{ER}) & \text{if } z_t = z_R \text{ and } z_{t-1} = z_E \\ \mathcal{N}(\mu_{RR}, \sigma^2_{RR}) & \text{if } z_t = z_R \text{ and } z_{t-1} = z_R \end{cases}. \tag{8}$$

Importantly, despite the shocks coming from multiple state-dependent distributions with different means and variances, this set-up restricts both the skewness and excess kurtosis of productivity growth in both aggregate states to be 0.

### *D. Firms*

I make the standard assumption of there being a continuum of measure 1 of firms that can be aggregated into a single representative firm. This representative firm demands capital and labor from the households and produces output according to the Cobb-Douglas production function

$$Y_t = F(K_t, L_t) = K_t^{\alpha} L_t^{1-\alpha}, \tag{9}$$

where $Y_t$ is output, $K_t$ is capital, $L_t$ is effective labor (a function of a household's productivity and labor choice), and $\alpha$ is capital's share of output. The firm maximizes profits and pays capital and labor their respective marginal products.

### *E. Government*

The government seeks to solve a Ramsey problem. It maximizes a utilitarian social welfare function,

$$\Omega = \max_{\{\tau_t\}_0^\infty} \mathbf{E}_0 \int_0^1 U_i \, di, \tag{10}$$

by setting the path of the tax progressivity parameter. This path is set *ex ante* such that $\tau_t$ is exclusively dependent on $z_t$: for both possible states $z_E$ and $z_R$, the government chooses $\tau_E$ and $\tau_R$ such that

$$\tau_t = \begin{cases} \tau_E & \text{if } z_t = z_E \\ \tau_R & \text{if } z_t = z_R \end{cases}, \tag{11}$$

Therefore, the government is constrained to not be able to actively change its tax policy from one period to the next if the aggregate state of the economy has not changed. Full commitment is assumed and the households know this.

While other papers in the literature focus on a single value for $\tau$ to act as an automatic stabilizer as in McKay and Reis (2021) or allow $\tau$ to change continuously as in Zoi (2020), I stay between these two extremes; restricting $\tau$ to one value ignores the government's ability to engage in any active fiscal policy, while allowing $\tau$ to change every period ignores the government's institutional frictions preventing it from enacting certain policies or enacting them too quickly. My approach takes both these issues into account, and allows for better analysis of the interactions between income tax progressivity and cyclical income risk.

The government is also subject to the following budget constraint:

$$G_t = \int_0^1 y_{it}^0 - y_{it} \, di \tag{12}$$

$$= \int_0^1 W_t e_{it} h_{it} - \lambda_t (W_t e_{it} h_{it})^{1-\tau_t} \, di. \tag{13}$$

The left-hand side is exogenous government spending. I assume it has no productive nor utility purposes to focus more on the redistributive aspects of progressive income taxation. This spending is a constant fraction $g$ of total output $Y_t$. The right-hand side is (net, after redistribution) revenue. To focus on the role of progressivity, I follow McKay and Reis (2021) and similar papers and ignore the possibility of government borrowing.[9] The government adjusts $\lambda_t$, the parameter governing the level (not progressivity) of the income tax, each period residually to make the budget constraint hold. This leaves $\tau_E$ and $\tau_R$ as the government's only choice variables.

### *F. Market Clearing*

There are three markets in this economy that must clear at all times. The first is the rental market for capital. The stock of capital used in production must equal the aggregate stock of capital held by households:

$$K_t = \int_0^1 k_{it}\ di. \tag{14}$$

The second market is the labor market. The amount of effective labor used in production must equal the aggregate product of household productivity and household labor supply:

$$L_t = \int_0^1 e_{it} h_{it}\ di. \tag{15}$$

This leaves the market for the consumption good, which clears when

$$Y_t = C_t + K_{t+1} - (1-\delta)K_t + G_t, \tag{16}$$

where $K_{t+1} - (1-\delta)K_t$ is net investment.

### *G. Equilibrium*

To solve dynamic equilibrium models numerically, it is useful to use recursive methods. Using recursive language involves expressing behavior and prices as a function of individual and aggregate state variables instead of time.

The recursive value function of a household for a given tax policy $\tau(z)$ is

$$V(e,k,z,D) = \max_{c,\ k',\ h} \Big\{ u(c,h) + \beta \mathbf{E}\Big[ V(e',k',z',D') | e, z \Big] \Big\} \tag{17}$$

[9]This is why households only have access to capital and not bonds.

subject to

$$c + k' = \lambda(Weh)^{1-\tau} + (1 + R - \delta)k, \tag{18}$$

$$k' \geq \bar{k}, \tag{19}$$

$$D' = \Gamma(D, z, z') \tag{20}$$

where $\Gamma$ denotes a transition operator that defines a law of motion for the cross-sectional distribution of households, $D(e, a)$.

The *recursive competitive equilibrium* for a given tax policy $\tau(z)$ consists of the value function $V(a, e, z, D)$, policy functions $c(a, e, z, D)$, $k'(a, e, z, D)$, and $h(a, e, z, D)$, aggregate capital, effective labor, and government spending $K(z, D)$, $L(z, D)$, $G(z, D)$, input prices $W(z, D)$ and $R(z, D)$, taxation level $\lambda(z, D)$ and a law of motion for the distribution $\Gamma(D, z)$ such that:

1) given $W$, $R$, $\lambda$, and $\tau$, the value function $V$ and the individual policy functions $c$, $k'$, and $h$ solve the households' decision problem.
2) the prices of the production inputs equal their marginal productivities, $W = F_L(K, L)$ and $R = F_K(K, L)$.
3) the input markets clear, $K = \int k \, dD$ and $L = \int eh \, dD$.
4) the goods market clears and the aggregate resource constraint is satisfied, $Y = C + K^{'} - (1 - \delta)K + G$.
5) the government balance its budget, $gY = \int Weh - \lambda(Weh)^{1-\tau} \, dD$.
6) Individual and aggregate behaviors are consistent, $D' = \Gamma(D, z, z')$.

Using this definition of equilibrium, I can define what optimal policy means. In terms of the household value function, the government's social welfare function is

$$\Omega = \max_{\tau(z)} \mathbf{E} \int V(a, e, z, D) dD. \tag{21}$$

The *optimal recursive competitive equilibrium* (the Ramsey plan) for the optimal tax policy $\tau^*(z)$ consists of the social welfare function $\Omega$ and the components of the recursive competitive equilibrium for a given tax policy such that:

1) Given all the other components, $\tau^*(z)$ solves $\Omega$.
2) $\tau^*(z)$ and all the other components are a recursive competitive equilibrium for $\tau^*(z)$.

## III. Quantitative Analysis

### A. Calibration

The unit of time is one year.[10] All data and targeted moments are for post-war US. Table 1 provides a list of all parameters values used for the analysis of the model. The parameter values come from a mix of internal and external calibration and are set to match moments of the model's steady-state.[11]

Table 1—Calibrated Parameters

| | Interpretation | Value | Target / Source |
|---|---|---|---|
| | *Preferences* | | |
| $\beta$ | Discount factor | 0.945/0.955 | Capital-output ratio of 2.7 |
| $\theta$ | Relative risk aversion scaler | 2.170/2.177 | Mean relative risk aversion of 3 |
| $\psi$ | Labor disutility weight | 0.919/0.798 | Normalize $L$ to 1 |
| $\varphi$ | Adjusted inverse Frisch elasticity | 1.442 | Inverse Frisch elasticity of 2 |
| | *Household Savings Choice* | | |
| $\underline{k}$ | Borrowing constraint | 0 | Standard |
| $\delta$ | Capital depreciation | 0.094 | Investment-output ratio of 0.255 |
| | *Firms* | | |
| $\alpha$ | Capital share | 0.36 | Standard |
| | *Government* | | |
| $g$ | Government fraction of output | 0.18 | Mean Federal Outlay % of GDP |
| $\tau$ | Income tax progressivity | 0.186 | HSV (2020) |
| | *Aggregate State Transitions* | | |
| $\Pi_{EE}$ | Prob. of staying in an expansion | 0.771 | NBER business cycle dating |
| $\Pi_{RR}$ | Prob. of staying in a recession | 0.421 | NBER business cycle dating |

**Notes:** Some parameters have two values because they are affected by the distributions of the individual productivity innovations: the left number is when the innovations come from a Johnson's $S_U$ distribution and the right is when innovations come from a normal distribution. The *Target/Source* numbers are the values in steady-state. *Standard* means it's a commonly used value and near the middle of the range of used values. For the business cycle dating, I follow the approach of Guvenen, Ozkan, and Song (2014) to be consistent with their estimated moments.

The discount factor $\beta$ is set to match a capital-output ratio of 2.7. The relative risk aversion controlling parameter $\theta$ is set to match an average relative risk aversion of 3, a commonly assumed value. The labor disutility weight $\psi$ is set to normalize the aggregate effective labor $L$ to 1. The calibration of these three parameters depends on the household productivity process. When shocks come from a Johnson's $S_U$ distribution, the parameter values are 0.945, 2.170, and 0.919, respectively. When they come from a normal distribution, they are 0.955, 2.177, and 0.798. While

[10]A shorter unit of time, such as one quarter, would be ideal. However, because the amount of money a household owes the government is based on total income earned over the entire year, this is not feasible without significant alteration to the tax function and household problem.

[11]Because of the Markov process nature of the aggregate shocks, there is not a steady-state in the usual sense. This is discussed later in section III.B.

the risk aversion values are very close, the discount factor and labor disutility weight are significantly different.

The differences in these two parameter values imply two things. First, considering $\beta$, *ceteris paribus*, an economy with skewed and fat-tailed productivity growth innovations will cause households to save more on average than in an economy with normal innovations. This makes sense, since increased volatility of productivity in an economy with incomplete markets leads to more capital accumulation due to the desire for precautionary savings. Here, the increase in precautionary savings comes not from increased variance, but specifically from increased tail risk.

Similarly for $\psi$, *ceteris paribus*, an economy with skewed and fat-tailed productivity growth innovations will cause households to work more on average. This again reflects the household's desire to earn more income to be more self-insured in the case of a significant negative shock to productivity.

Regardless of the productivity process, $\theta$ does not equal the assumed average risk aversion of 3. The reason these numbers are not the same is due to how consumption and labor are linked in the household utility function.[12] A household's degree of relative risk aversion $R(c)$ is given by

$$R(c) \equiv -\frac{c u_{cc}}{u_c} = \theta \frac{c}{c - \psi \frac{h^{1+\varphi}}{1+\varphi}}. \tag{22}$$

Because of the remaining disutility of labor term, $R(c) \neq \theta$ as it does for some other kinds of preferences. Instead, equation 22 implies that households with a higher consumption-labor ratio are less risk averse.

Similarly, the progressive tax-adjusted inverse Frisch elasticity is set to 1.442 to match an inverse Frish elasticity of 2. This discrepancy in values comes from the effect of the log-linear tax function on a household's labor supply,

$$h = \left( \frac{(1-\tau)\lambda(We)^{1-\tau}}{\psi} \right)^{\frac{1}{\varphi+\tau}}, \tag{23}$$

because the Frisch elasticity, $\frac{\partial \ln h}{\partial \ln w}$, is the exponent on the wage rate, $\frac{1-\tau}{\varphi+\tau}$, which depends on the the degree of progressivity.

The capital depreciation rate $\delta$ is set to 0.094 to match an investment-output ratio of 0.255. I assume a borrowing constraint of $\underline{k} = 0$ and a capital share of output of $\alpha = 0.36$. The government fraction of GDP $g$ is set equal to the average federal net outlays as a percent of GDP, 0.18. The

[12]The cross-derivative of utility with respect to consumption and labor is not zero.

income tax progressivity controlling parameter $\tau$ is taken from Heathcote, Storesletten, and Violante (2020), who finds an estimate of 0.186 by regressing the log-difference between pre-government income and taxes minus transfers on log-disposable income.

Calibrating the aggregate state transitions involves calculating the probability that the economy will be in an expansion/recession next period given that it is in an expansion/recession this period. Because the unit of time is one year, I must take a stand on what years are expansionary and recessionary. I use the exact same classification system as Guvenen, Ozkan, and Song (2014) to be consistent with their estimated moments of individual income growth. Within their sample extending from 1977 to 2011, the years 1980-1983, 1991-1992, 2001-2002, and 2008-2010 are considered recessionary. With these ranges and the properties of a Markov process, I calculate that the probability $\Pi_{EE}$ of being in an expansion next year given an expansion this year is 0.771: the probability $\Pi_{ER}$ of being in a recession next year given an expansion this year is $1 - \Pi_{EE} = 0.229$: the probability $\Pi_{RR}$ of being in a recession next year given a recession this year is 0.421: the probability $\Pi_{RE}$ of being in an expansion next year given a recession this year is $1 - \Pi_{RR} = 0.579$.

The final part of the calibration strategy is the individual productivity process. The parameter values are in Table 6 in Appendix V.B. The parameters are calibrated using the method of simulated moments (MSM) to match the average mean, standard deviation, standardized skewness, and standardized excess kurtosis for individual productivity growth in expansions and recessions. Table 2 reports the statistics from the data and from simulations of the productivity process using Johnson's $S_U$ and normal shocks after finding the best parameter values. While normal shocks can perfectly match the standard deviations, as expected it is unable to match the skewness and kurtosis. The Johnson's $S_U$ shocks have no problem matching the higher moments. While both struggle to match the means, this is less important than adequately matching the other moments because productivity will be normalized to have an unconditional mean of 1.

### *B. Computation*

Models with heterogeneous agents, incomplete markets, and aggregate fluctuations are notoriously difficult to solve, due to the entire distribution of households being an infinite-dimensional state variable that each household and the government must track. However, recent advancements in numerical methods have allowed economists to solve these types of models more quickly and accurately. I use the now standard techniques of Carroll (2006) and Young (2010) to solve for the model's steady state. From there I adapt the techniques of Boppart, Krusell, and Mitman (2018) and Auclert et al. (2021) to solve for the model's transition paths caused by aggregate fluctuations.

Table 2—Targeted Moments of Income Growth

| Moment | Data | $S_U$ Model | Normal Model |
|---|---|---|---|
| $\mu_E$ | 0.033 | 0.007 | 0.007 |
| $\sigma_E$ | 0.510 | 0.536 | 0.510 |
| $s_E$ | -0.830 | -0.819 | 0.004 |
| $\kappa_E$ | 11.930 | 12.020 | -0.073 |
| $\mu_R$ | -0.008 | -0.018 | -0.018 |
| $\sigma_R$ | 0.510 | 0.497 | 0.510 |
| $s_R$ | -1.680 | -1.6767 | 0.002 |
| $\kappa_R$ | 11.930 | 11.993 | -0.215 |

**Notes:** This table reports statistics for the mean, standard deviation, standardized skewness, and standardized excess kurtosis of productivity growth in expansions and recessions. The moments estimated from data come from Guvenen, Ozkan, and Song (2014) and Guvenen et al. (2021). The moments from the models are the estimates from simulations after solving for the productivity parameters.

To computational solve the model, I must discretize the household productivity process. To do this, I fix a discrete grid of values for $e$ and use the method of Tauchen (1986) to calculate the matrices of individual transition probabilities for each aggregate transition, $\Pi^e_{EE}$, $\Pi^e_{ER}$, $\Pi^e_{RE}$, and $\Pi^e_{RR}$, where

$$\Pi^e_{zz'} = \Pr(e^{'}|e, z, z^{'}). \tag{24}$$

for each combination of $e$ and $e^{'}$ using the discrete grid.

Because the aggregate shocks come from a Markov process, the standard notion of steady-state doesn't apply to my model. Instead, the relevant steady state is when the transition matrix for household productivity does not depends on $z$. This aggregate state-independent transition matrix is given by the weighted average of the state-dependent transition matrices,

$$\Pi^e_{ss} = \pi_E(\Pi_{EE}\Pi^e_{EE} + \Pi_{ER}\Pi^e_{ER}) + \pi_R(\Pi_{RR}\Pi^e_{RR} + \Pi_{RE}\Pi^e_{RE}), \tag{25}$$

where $\pi_E$ and $\pi_R$ are the unconditional probabilities for expansion and recession, respectively.[13] Similarly, the steady-state degree of progressivity is

$$\tau_{ss} = \pi_E\tau_E + \pi_R\tau_R, \tag{26}$$

From there I utilize a version of the endogenous grid method of Carroll (2006) to obtain the stationary value function $V_{ss}$ and consumption $c_{ss}$, savings $a^{'}_{ss}$, and labor $h_{ss}$ policy functions of

[13]If the length of the simulation path discussed later, $T$, is not large enough, one may want to use the aggregate transition probabilities and unconditional probabilities implied by the simulation instead of the true values. This can help avoid numerical instability when calculating the transition paths.

the households and use said policy functions and the lottery method of Young (2010) to obtain the stationary distribution, $D_{ss}$.

After the stationary distribution is calculated, I can look at how a change in aggregate productivity affects every other variable, both at the time of the change and all future periods. This is usually formulated as an MIT shock, such as in the recent advancements of Boppart, Krusell, and Mitman (2018) and Auclert et al. (2021). They suppose that an exogenous one-time shock to a aggregate variable (such as TFP or the interest rate) occurs while the economy is at its steady state such that $D_1 = D_{ss}$ and the economy will return to its steady state after $T$ periods so that $V_{T+1} = V_{ss}$.[14] To do this, the value and policy functions are iterated backward to get the paths of these functions. Once the path of the savings policy function is known, a path for the distribution is obtained by forward iteration. The paths of all relevant aggregate variables can be calculated from the paths of the policy functions and distributions, and certain aggregates are manipulated to ensure all market clearing conditions hold. If any markets in any period are not cleared within some tolerance level, then the path(s) of one or more endogenous variables are updated and the process is repeated until convergence.

Their technique as is does not work for my model, as 1) my aggregate fluctuations cannot be expressed well as an MIT shock, and 2) there is no true steady state at which to start nor end. I extend their idea by having my sequence for the exogenous aggregate variable not be an MIT shock. Instead of an aggregate variable experiencing a one-time increase that slowly dissipates back to steady state, I create a simulation of my specified aggregate Markov process for $T+1$ periods such that all individual and aggregate variables in periods 1 and $T+1$ are the same. This eliminates the need for a steady state at which to start or end, because it is as if periods 1 through $T$ are repeated indefinitely.

To accomplish this, I first do one iteration of the backward and forward iteration of $V_t$ and $D_t$, respectively, using the steady state created by the transition matrix in equation 25. This will create $V_1 \neq V_{T+1}$ and $D_1 \neq D_{T+1}$. The trick is to update $V_{T+1}$ to $V_1$ and $D_1$ to $D_{T+1}$ and continue 1) performing the backward and forward iteration of the value function and distribution paths, and 2) updating the terminal value function and initial distribution. I stop repeating these steps when a measure of convergence is reached.

One downside of the Boppart, Krusell, and Mitman (2018) and Auclert et al. (2021) techniques is that the path of aggregate variable is deterministic and known to the households with perfect

[14]Following Auclert et al. (2021), I set $T = 300$. While they suggest this value for a model with quarterly frequency, I still use 300 instead of the implied 75 for annual frequency to help prevent the randomness of the simulated aggregate path from significantly affecting results.

foresight; the households know when each expansion and recession will start and end. However, to my knowledge, a solution technique for this type of heterogeneous-agent model with aggregate fluctuations that maintains uncertainty at the aggregate level is currently not available. Furthermore, models utilizing MIT shocks (intentionally) face this same issue. It is important to note that while there is no uncertainty regarding the aggregate state, uncertainty is maintained with respect to the individual productivities. While a household may know that an expansion is about to turn into a recession next period, they still don't know which idiosyncratic state they will have next period. They only know the transition probabilities.

Using my new solution method, I can solve the model's steady state and transition paths for a given government policy $\tau(z)$ and calculate the fluctuating economy's *ex ante* expected social welfare:

$$\Omega = \frac{1}{T} \sum_{t=1}^{T} \sum_{e,\ k} D_{ekt} V_{ekt}. \tag{27}$$

Here I use the phrase *ex ante* with some abuse of terminology. Strictly speaking, *ex ante* usually means all decisions and evaluations are made before any events occur, i.e. before anything is known about the first time period. However, my solution technique requires the path of all variables to be recursive after $T$ periods, so there's no true "first period". Therefore, I mean *ex ante* in the sense that the government will optimize social welfare assuming that there is an equal chance that any of the $T$ periods could be the first period.[15]

As social welfare is an ordinal measure, it is impossible to directly compare the value of social welfare in a baseline economy to its value in the corresponding restricted and unrestricted economies. I alleviate this issue by converting expected life-time utility into more interpretable units by calculating the consumption equivalent variation (CEV), $\omega$. The CEV in this context is how much every household in every time period would need to receive, in percentage terms of consumption, to be *ex ante* indifferent between living in the baseline economy or in the alternative policy economy. $\omega$ solves

$$\sum_{t=1}^{T} \sum_{e,\ k} D_{ekt}^{base} V_{ekt}^{base}((1+\omega)c_{ekt}, h_{ekt}) = \sum_{t=1}^{T} \sum_{e,\ k} D_{ekt}^{alt} V_{ekt}^{alt}(c_{ekt}, h_{ekt}) \tag{28}$$

where, with some abuse of notation, I've rewritten the value function to have the consumption

[15]A motivation for this interpretation of *ex ante* comes from extending Rawls (1971)'s "veil of ignorance" idea to mean making decisions not only based off not knowing what productivity and capital a household will have (i.e. cross-sectional ignorance), but also not knowing from what moment in time a household will start (i.e. inter-temporal ignorance).

and labor policy functions as its inputs. This measure of welfare adjusts for risk, intertemporal substitution, and mean-reversion in ability.

### C. *Results*

I solve the model six times, three times per productivity process. For each productivity process, the first scenario is the baseline economy; the progressivity parameters are set to $\tau_E = \tau_R = 0.186$, which is the progressivity value used in calibration, and the government is not maximizing social welfare. The second scenario is where the government's policy is restricted to be state-independent, so $\tau_E = \tau_R$, and chooses one progressivity value to maximize welfare. The third scenario is where the government's policy is relaxed to be state-dependent, so $\tau_E \neq \tau_R$, and chooses both values to maximize welfare. The results of these scenarios are in Table 3.

Table 3—Optimal Progressivity Values and Welfare Results

| | Baseline | Restricted Policy | Unrestricted Policy |
|---|---|---|---|
| Optimal Progressivity | | | |
| Johnson's $S_U$ | 0.186 | 0.111 | (0.125, 0.072) |
| Normal | 0.186 | 0.006 | (0.011, -0.004) |
| CEV Welfare Change | | | |
| Johnson's $S_U$ | NA | 0.8 | 0.9 |
| Normal | NA | 2.8 | 2.9 |

**Notes:** This table reports, for each policy-distribution pair, the values(s) for the optimal progressivity parameter(s) and the resulting consumption equivariant variation welfare gains. The welfare gains are expressed in percentages. The baseline column is for reference.

I find that the current degree of income tax progressivity in the US is too high regardless of policy restriction and innovation distribution. There are positive welfare gains by decreasing progressivity in every scenario. However, the difference in welfare gains between policy restrictions is much smaller than the difference in gains between innovation distributions. I consider the former first.

When comparing restricted and unrestricted policy, I find significant differences in progressivity values. When unrestricted, it is optimal for the government to set progressivity higher in expansions than in recessions. When restricted, the government, expectantly, chooses to set progressivity between the two unrestricted values. However, the difference in welfare gains is a relatively small 0.1%, regardless of the innovation distribution. Together, these findings imply the curvature of the social welfare function $\Omega(\tau_E, \tau_R)$ is fairly flat in the neighborhood of the restricted and unrestricted policy progressivity values but relatively more curved near the baseline. This provides evidence in support of treating the progressivity of income taxes as an automatic stabilizer like in McKay and

Reis (2021).

Next, the choice of innovation distribution significantly impacts the optimal degree of progressivity and the associated welfare change. In the presence of skewed and fat-tailed productivity growth, the government finds it optimal to lower the income tax progressivity parameter by about 40%, from 0.186 to 0.111 (for the restricted policy). In contrast, if the higher moments are shut down, then the optimal parameter for a model using normal innovations is approximately 0.006. With a progressivity parameter so close to zero, the income tax function is essentially a flat tax for all but the few extremely productive households. This result is not surprising, since the variance of the idiosyncratic shocks is approximately constant over time. It is a well-known theoretical result from similar classes of models that if only the mean of a distribution is shifting, then a flat tax rate is welfare-maximizing. This theoretical result applies here.

Table 4 shows the responses of several household variables to changes in tax progressivity for different households with low, medium, and high productivity. Specifically, it reports the average elasticity of a variable; the average percent change of the variable, in either expansions or recessions, due to a 1% change in one of the progressivity parameters. The household variables shown are value functions; consumption, savings, and labor policy functions; and pre-tax and post-tax incomes. There are several things to note here. First, I look only at the elasticities for the more interesting case of the economy with non-normal shocks. Second, these are point elasticities: they are evaluated at the restricted optimal policy of $\tau_E = \tau_R = 0.111$ to see why the unrestricted policy is to have higher progressivity in expansions. Third, the elasticities reported are for for households whose personal amount of capital equals the overall average amount of capital in the economy.

Several findings emerge. Every elasticity is less than one in absolute value, indicating these variables are inelastic to progressivity changes. This is not surprising, as the progressivity parameters are at the point in the parameter space where social welfare is maximized with the restraint that the parameters must be equal. Looking at the elasticities for the value functions, an increase in expansion progressivity increases the value functions in *both* expansions and recessions more than an increase in recession progressivity (for all but the most productive households). This is the direct reason why the unrestricted optimal policy is to have higher progressivity in expansions. Furthermore, increasing $\tau_E$ causes a bigger increase in $V_R$ than in $V_E$, while increasing $\tau_R$ causes a bigger increase in $V_E$ than in $V_R$. Counterintuitively, increasing redistribution in expansions does a better job of improving the situation of lower-productivity households in recessions.

The resolution to this paradox comes from how increasing progressivity affects both the pre-tax income and post-tax income of the low-productivity households (the ones who impact social welfare most through their higher marginal utilities). Increasing $\tau_E$ increases both incomes for both aggregate states, while increasing $\tau_R$ decreases pre-tax income and post-tax income in recessions. This is interesting because the rationale for having more redistribution in recessions is to increase post-tax income of the poor at the times when they are worst off, regardless of their pre-tax income decreasing due to general equilibrium (GE) effects from changes in aggregate variables. I find that this rationale doesn't hold; increasing redistribution in recessions leads to the opposite of the desired result. This comes from the aforementioned GE effects, as an increase in $\tau_E$ changes aggregate capital $K_R$ and output $Y_R$ in such a way that both the wage rate $W_R$ and tax level $\lambda_R$ increase, allowing higher post-tax income, *despite* the low-productivity households working *less*.

Table 4—Household Variables

| | Low $e$ | | Medium $e$ | | High $e$ | |
|---|---|---|---|---|---|---|
| | $\tau_E$ | $\tau_R$ | $\tau_E$ | $\tau_R$ | $\tau_E$ | $\tau_R$ |
| $V_E$ | 0.007 | 0.004 | 0.006 | 0.003 | -0.009 | 0.000 |
| $V_R$ | 0.010 | 0.002 | 0.007 | 0.002 | 0.000 | -0.009 |
| $c_E$ | 0.012 | 0.008 | -0.009 | -0.003 | -0.793 | -0.024 |
| $c_R$ | 0.018 | 0.003 | -0.010 | -0.021 | -0.071 | -0.817 |
| $k'_E$ | 0.002 | 0.003 | 0.011 | 0.001 | -0.542 | -0.018 |
| $k'_R$ | 0.008 | -0.002 | 0.003 | 0.003 | -0.058 | -0.570 |
| $h_E$ | 0.619 | -0.010 | -0.031 | -0.010 | -0.356 | -0.010 |
| $h_R$ | -0.031 | 0.608 | -0.031 | -0.042 | -0.031 | -0.367 |
| $\lambda_E y_E^{1-\tau_E}$ | 0.026 | 0.026 | 0.036 | 0.004 | -0.727 | -0.023 |
| $\lambda_R y_R^{1-\tau_R}$ | 0.079 | -0.019 | 0.006 | -0.007 | -0.072 | -0.756 |
| $y_E$ | 0.025 | 0.026 | -0.010 | 0.001 | -0.366 | -0.024 |
| $y_R$ | 0.079 | -0.019 | -0.004 | -0.030 | -0.073 | -0.353 |

**Notes:** This table reports sets of point elasticities for the minimum, mode, and maximum values of the grid for the discretized productivity process. The numbers represent the percent change of a household-level variable averaged over expansion periods or recession periods due to a 1% change in the progressivity parameter for expansions or recessions. These elasticities are calculated for the economy with the skewed productivity process and evaluated at the restricted optimal policy for households with an amount of capital equal to the average amount of capital.

Before concluding, I will briefly discuss the aggregate outcomes of the two productivity models. Table 5 reports the aggregate outcomes for consumption, capital, effective labor, output, labor, the rental rate of capital, the wage rate of labor, and the level of taxation. First, regardless of the productivity innovation distribution and whether the government's policy is restricted or unrestricted, lowering income tax progressivity from the current US value leads to higher aggregate consumption, capital, and output. This increase in production also has the general equilibrium effect of letting the government lower the average *level* of taxation, despite the increased output causing required government expenditure to rise. All this provides evidence that the US may on average be on the undesirable right side of the Laffer curve, and cutting taxes can increase government revenue.

Table 5—Aggregate Variables

| | Johnson's $S_U$ Distribution | | | Normal Distribution | | |
|---|---|---|---|---|---|---|
| | Baseline | Restricted | Unrestricted | Baseline | Restricted | Unrestricted |
| $C_E$ | 0.973 | 9.75 | 8.90 | 0.975 | 18.38 | 18.10 |
| $C_R$ | 0.921 | 9.10 | 11.08 | 0.963 | 18.12 | 18.68 |
| $K_E$ | 4.618 | 14.83 | 14.65 | 4.706 | 24.77 | 24.77 |
| $K_R$ | 4.289 | 15.33 | 15.50 | 4.444 | 25.65 | 25.72 |
| $L_E$ | 1.042 | 9.26 | 8.06 | 1.001 | 17.77 | 17.43 |
| $L_R$ | 0.866 | 8.88 | 11.89 | 0.974 | 17.88 | 18.55 |
| $Y_E$ | 1.777 | 11.24 | 10.40 | 1.743 | 20.25 | 20.03 |
| $Y_R$ | 1.516 | 11.14 | 13.15 | 1.670 | 20.63 | 21.10 |
| $H_E$ | 0.815 | 5.52 | 5.07 | 0.881 | 12.70 | 12.53 |
| $H_R$ | 0.779 | 6.65 | 8.34 | 0.866 | 13.15 | 13.50 |
| $R_E$ | 0.139 | -3.14 | -3.72 | 0.133 | -3.65 | -3.82 |
| $R_R$ | 0.128 | -3.70 | -2.13 | 0.135 | -4.04 | -3.72 |
| $W_E$ | 1.098 | 1.81 | 2.15 | 1.122 | 2.10 | 2.21 |
| $W_R$ | 1.174 | 2.18 | 1.26 | 1.122 | 2.33 | 2.14 |
| $\lambda_E$ | 0.796 | -1.78 | -1.10 | 0.763 | -5.49 | -5.21 |
| $\lambda_R$ | 0.742 | 0.25 | -0.75 | 0.754 | -4.36 | -4.89 |

**Notes:** The two baseline columns report the value of the aggregate variables averaged over expansion periods or recession periods. The values are relative to the normalization of $L = 1$ in steady-state as mentioned in section III.A. The other four columns report the variables' percent changes from their respective baseline for each policy-distribution pair. $H$ is the aggregate household labor. This is distinct from the aggregate effective labor $L$, as $H$ doesn't incorporate household productivity.

Second, while also present when assuming a normal distribution, the aggregate variable changes show distinct differences between the restricted and unrestricted policy for the Johnson's $S_U$. Because progressivity is substantially lower in recessions than in expansions with the unrestricted policy, I observe consumption, capital, effective labor, output, and labor having a larger percent increase in recessions than in expansions.

Finally, in all scenarios, due to the large increase in the supply of capital—the most changed aggregate variable—the rental rate falls. However, even though the supply of labor increases, the capital-labor ratio also increases, causing the wage rate to rise.

## IV. Conclusion

This paper asks how a utilitarian government that puts equal weight on all households in the economy should vary income tax progressivity between expansions and recessions. My findings show that taxes should be more progressive in expansions than in recessions, but the difference is minor compared to the overall degree of progressivity. The overall progressivity should be approximately 40% lower than the historical average in the US.

This optimal level of progressivity heavily depends on the distribution of shocks to individual productivity growth. When shocks come from normal distributions—the standard assumption in most business cycle research—the optimal policy is to have the income tax function be close to a flat, proportional tax rate. This is a common result from many papers in related literature, and implies significant welfare gains given the US's current degree of progressivity. However, if shocks to productivity growth are left-skewed and fat-tailed, i.e. negative shocks are uncommon but likely to be very large while positive shocks are common but likely to be small, then the optimal policy is still to lower progressivity but not to the extreme of having income taxes be close to proportional. By misspecifying the model to have normal innovations, welfare gains are tripled from approximately 1% to 3%. Policy recommendations by research that ignores the higher moments of income growth innovations may significantly overstate the policy's effects and desirability.

## V. Appendix

### A. *Johnson's SU Distribution*

The Johnson's $S_U$-distribution is a four-parameter probability distribution first investigated by Johnson (1949) and used in various economics and finance papers such as Cayton and Mapa (2015), Kar et al. (2024), and Choi and Min (2025). The choice of using Johnson's $S_U$ in this paper is for convenience but is somewhat arbitrary, as there exist other four-parameter distributions that can also be used to generate skewness and kurtosis.

By adequately transforming the normal distribution, the distribution can allow for any amount of skewness and positive excess kurtosis.[16] If

$$X = \zeta \sinh\left(\frac{Z-\gamma}{\eta}\right) + \xi \tag{29}$$

and $Z \sim \mathcal{N}(0,1)$, then $X \sim S_U(\xi, \zeta, \gamma, \eta)$, where $\xi$ and $\gamma$ are any real numbers and $\zeta$ and $\eta$ are real numbers greater than zero. $\lambda$ and $\delta$ are the standard symbols used, not $\zeta$ and $\eta$, respectively, but those letters were already used in this paper to represent their standard meanings in similar economic models. $\xi$ controls the mean, $\zeta$ controls the variance, $\gamma$ controls the skewness, and $\eta$ controls the kurtosis. Closed-form expressions for the parameters as functions of the four moments are not available and numerical methods are required to calculate them.

The distribution's probability density function (PDF) is given by

$$p(x|\xi,\zeta,\gamma,\eta) = \frac{\eta}{\gamma\sqrt{2\pi}} \frac{1}{\sqrt{1+\left(\frac{x-\xi}{\zeta}\right)^2}} e^{-\frac{1}{2}\left(\gamma+\delta \sinh^{-1}\left(\frac{x-\xi}{\zeta}\right)\right)^2} \tag{30}$$

with a support of $-\infty$ to $\infty$.

[16]Johnson's $S_B$ (different from $S_U$) distribution is used for negative excess kurtosis.

### B. Productivity Process Discretization

Table 6—Individual Productivity Process Parameters

| $S_U$ Distributions | | Normal Distributions | |
|---|---|---|---|
| Parameters | Value | Parameters | Value |
| $\rho$ | 0.5 | $\rho$ | 0.5 |
| $\xi_{EE}$ | -0.997 | $\mu_{EE}$ | -0.192 |
| $\zeta_{EE}$ | 0.214 | $\sigma_{EE}$ | 0.395 |
| $\gamma_{EE}$ | 0.672 | $\mu_{RE}$ | -0.317 |
| $\eta_{EE}$ | 1.050 | $\sigma_{RE}$ | 0.477 |
| $\xi_{RE}$ | -1.829 | $\mu_{ER}$ | -0.325 |
| $\zeta_{RE}$ | 2.054 | $\sigma_{ER}$ | 0.400 |
| $\gamma_{RE}$ | -1.462 | $\mu_{RR}$ | -0.138 |
| $\eta_{RE}$ | 4.959 | $\sigma_{RR}$ | 0.415 |
| $\xi_{ER}$ | -1.663 | | |
| $\zeta_{ER}$ | $6\times10^{-5}$ | | |
| $\gamma_{ER}$ | 0.920 | | |
| $\eta_{ER}$ | 0.376 | | |
| $\xi_{RR}$ | -0.611 | | |
| $\zeta_{RR}$ | 0.089 | | |
| $\gamma_{RR}$ | 1.391 | | |
| $\eta_{RR}$ | 0.972 | | |

**Notes:** $\rho$ is set to exactly match the Guvenen, Ozkan, and Song (2014)'s estimate of income growth autocorrelation, -0.25. The other parameters are found using an adaptive differential evolution global optimization algorithm to solve the moment matching problem described in section III.A.

## REFERENCES

**Aiyagari, S. Rau.** 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *The Quarterly Journal of Economics*, 109(3): 37–56.

**Angelopoulos, Konstantinos, Stylianos Asimakopoulos, and James Malley.** 2019. "The Optimal Distribution of the Tax Burden over the Business Cycle." *Macroeconomic Dynamics*, 23(6): 2298–2337.

**Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub.** 2021. "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models." *Econometrica*, 89(5): 2375–2408.

**Bakis, Ozan, Baris Kaymak, and Markus Poschke.** 2015. "Transitional Dynamics and the Optimal Progressivity of Income Redistribution." *Review of Economic Dynamics*, 18(3): 679–693.

**Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J. Sargent.** 2021. "Inequality, Business Cycles, and Monetary-Fiscal Policy." *Econometrica*, 89(6): 2559–2599.

**Boar, Corina, and Virgiliu Midrigan.** 2021. "Efficient Redistribution." *Working Paper*.

**Boppart, Timo, Per Krusell, and Kurt Mitman.** 2018. "Exploiting MIT Shocks in Heterogeneous Agent Economics: The Impulse Response as a Numerical Derivative." *Journal of Economic Dynamics and Control*, 89: 68–92.

**Carroll, Christopher D.** 2006. "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems." *Economics Letters*, 91(3): 312–320.

**Cayton, Peter Julian, and Dennis Mapa.** 2015. "Time-Varying Conditional Johnson SU Density in Value-at-Risk Methodology." *Philippine Review of Economics*, 51(1): 23–44.

**Choi, Pilsun, and Insik Min.** 2025. "Estimating the Distribution of Net Worth and Disposable Income Using Johnson's SU Distribution." *Journal of Economic Theory and Econometrics*, 36(1): 1–14.

**Conesa, Juan Carlos, and Dirk Krueger.** 2006. "On the Optimal Progressivity of the Income Tax Code." *Journal of Monetary Economics*, 53(7): 1425–1450.

**Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger.** 2009. "Taxing Capital? Not a Bad Idea after All!" *American Economic Review*, 99(1): 25–48.

**Dyrda, Sebastian, and Marcelo Pedroni.** 2023. "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk." *The Review of Economics Studies*, 90(2): 744–780.

**Feldstein, Martin.** 1973. "On the Optimal Progressivity of the Income Tax." *Journal of Public Economics*, 2(4): 357–376.

**Ferrière, Axelle, Philipp Grubener, Gaston Navarro, and Oliko Vardishvili.** 2021. "Larger Transfers Financed with More Progressive Taxes? On the Optimal Design of Taxes and Transfers." *Working Paper*.

**Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song.** 2021. "What Do Data on Millions of U.S. Workers Reveal About Life-Cycle Dynamics?" *Econometrica*, 89(5): 2303–2339.

**Guvenen, Fatih, Serdar Ozkan, and Jae Song.** 2014. "The Nature of Countercyclical Income Risk." *Journal of Political Economy*, 122(3): 621–660.

**Hamilton, James D.** 1990. "Analysis of Time Series Subject to Changes in Regime." *Journal of Econometrics*, 45(1): 39–70.

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

**Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante.** 2020. "How Should Tax Progressivity Respond to Rising Income Inequality?" *Journal of the European Economic Association*, 18(6): 2715–2754.

**Johnson, N. L.** 1949. "Systems of Frequency Curves Generated by Methods of Translation." *Biometrika*, 36(1/2): 149–176.

**Kar, Tapan, Sesha Meka, Shiba Prasad Mohanty, and Kaveri Krishnan.** 2024. "Modelling Multivariate Return Volatilities by Independent Components Using Johnson SU Likelihood." *Global Business Review.*

**Krueger, Dirk, and Alexander Ludwig.** 2013. "Optimal Progressive Labor Income Taxation and Education Subsidies When Education Decisions and Intergenerational Transfers are Endogenous." *American Economic Review*, 103(3): 496–501.

**Krusell, Per, and Anthony A. Smith, Jr.** 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106(5): 867–896.

**McKay, Alisdair, and Ricardo Reis.** 2021. "Optimal Automatic Stabilizers." *Review of Economic Studies*, 88(5): 2375–2406.

**Rawls, John.** 1971. *A Theory of Justice.* Cambridge, MA: Belknap Press.

**Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron.** 2004. "Cyclical Dynamics in Idiosyncratic Labor Market Risk." *Journal of Political Economy*, 112(3): 695–717.

**Tauchen, George.** 1986. "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions." *Economics Letters*, 20(2): 177–181.

**Young, Eric R.** 2010. "Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell-Smith Algorithm and Non-Stochastic Simulations." *Journal of Economic Dynamics and Control*, 34(1): 36–41.

**Zoi, Sarah.** 2020. "Optimal Progressivity Over the Business Cycle." *Working Paper.*