TAXES ON LIFETIME INCOME: A GOOD IDEA?

Dirk Krueger
Chunzan Wu

Working Paper 33664
http://www.nber.org/papers/w33664

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2025, Revised March 2026

We wish to thank Jonathan Heathcote for useful feedback at an early stage of this paper and Marek Kapička and Vincenzo Quadrini as well as two anonymous referees for important comments that substantially improved the paper. All errors remain our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Taxes on Lifetime Income: A Good Idea?
Dirk Krueger and Chunzan Wu
NBER Working Paper No. 33664
April 2025, Revised March 2026
JEL No. E60, H20

ABSTRACT

In standard life-cycle models, household consumption and welfare more strongly depend on lifetime income, but most countries base income taxes on current income and use progressive taxes to reduce inequality and provide social insurance. Is lifetime income a better tax base for governments seeking to provide such social insurance and redistribution? To answer this question, we build a quantitative life-cycle model of heterogeneous households with idiosyncratic wage risks and endogenous labor supply, and calibrate it to the U.S. economy. We document that switching to a lifetime income tax leads to a more efficient distribution of hours worked over time and across states of the world. This benefit rises with tax progressivity under a lifetime income tax, whereas the opposite is true under an annual income tax. Consequently, the optimal lifetime income tax is more progressive and achieves larger ex-ante welfare for a cohort of households than the optimal annual income tax.

Dirk Krueger
University of Pennsylvania
Department of Economics
and CEPR
and also NBER
dkrueger@econ.upenn.edu

Chunzan Wu
Peking University
chunzan@sas.upenn.edu

1 Introduction

Most industrialized countries use a progressive income tax to redistribute resources from households with high market incomes to those with low market incomes, and to provide social insurance against idiosyncratic income fluctuations over time. Typically, the tax law bases the income tax solely on a household's income from the current year, independent of earnings in other years. However, a progressive income tax with rising marginal tax rates introduces labor supply distortions, relative to the efficient allocation that might compromise the potential welfare benefits from social insurance and redistribution. An annual progressive income tax is especially detrimental in this regard when households experience strong deterministic or stochastic variations in their labor productivity. In that case, efficiency, at least with preferences that are separable between consumption and labor, would dictate strong variation of labor supply following productivity over the life cycle. However, with a progressive tax system this variation in labor supply drives up the average tax rate for these households, relative to similar households facing more stable productivity profiles.

As already argued by Vickrey (1939), a tax on lifetime income naturally reduces such distortions and therefore may prove to be a more efficient tool for income redistribution and insurance. In the simplest case when there is no income risk and households can transfer resources freely across time through asset markets, lifetime income is a sufficient statistic for household welfare, and therefore, an ideal target for welfare redistribution. However, households are subject to income shocks over life cycle, may face potentially binding borrowing constraints and must make labor supply and consumption decisions before the uncertainty about their lifetime income is fully resolved. Realized lifetime income is then no longer a perfect measure of household welfare, and whether a lifetime income tax is still preferable under these circumstances is a quantitative question.

In this paper, we explore the positive and welfare properties of a lifetime income tax (LIT) system and contrast it with the status quo in which income taxes are based on annual earnings (AIT). First, we use a two-period model with endogenous labor supply and income risk in the second period to clarify the main trade-off between the two tax systems. We demonstrate theoretically that a lifetime income tax is more conducive to an efficient allocation of labor, both across time and across states of the world, but that it compromises the provision of consumption insurance later in the life cycle, relative to a tax on current income.[1]

To quantify this trade-off, we then construct a quantitative incomplete-markets life-cycle model featuring households that are ex ante heterogeneous with respect to initial labor pro-

---

[1] Our theoretical results on the allocation of labor are especially sharp when households are risk-neutral (Corollaries 1, 2, 4), and, therefore, consumption insurance considerations are absent.

ductivity and wealth as well as preferences for labor, and that are exposed to uninsurable idiosyncratic, non-Gaussian labor productivity shocks, inducing further ex-post heterogeneity.[2] Households make consumption-savings and labor supply decisions. The income tax policy is encoded in a nonlinear tax function that links household pre-tax earnings to their tax liability.[3] Under an annual income tax, both pre-tax earnings and tax liabilities pertain to the current year. In contrast, a lifetime income tax bases total tax liabilities on (discounted) lifetime earnings. We implement the lifetime income tax such that in every period households pay the increment in their lifetime tax liabilities due to the increase in their accumulated lifetime income resulting from their current earnings.[4]

The production side of our model is standard and consists of perfectly competitive firms with a constant-returns-to-scale technology that combine labor and capital to produce a final good that can be used for consumption and capital investment. In our benchmark analysis we assume a linear production technology, and thus factor prices are unaffected by tax reforms. This assumption (which effectively amounts to a partial equilibrium analysis) is relaxed in the robustness analysis, where we consider an alternative scenario in which factor prices are fully endogenous and determined by domestic market clearing.

The government, in addition to the labor income tax, the focus of our analysis, also collects revenues through consumption-, capital income-, and payroll taxes to fund expenditures on public goods, retirement benefits, and interest payments on government debt. The model is calibrated to mirror the U.S. economy between 1999 and 2017. Using this quantitative model, we then quantify the welfare gains from switching the tax base to lifetime income. We measure welfare as expected lifetime utility of a cohort born into the stationary equilibrium of the model. To insure comparability across tax systems we require that all systems collect

---

[2] In a previous version of this paper we modeled labor productivity as following a simple AR(1) process. We now follow recent advances in the literature and estimate and then implement a non-Gaussian *labor productivity (wage)* process of the form studied by Arellano, Blundell, and Bonhomme (2017), De Nardi, Fella, and Paz-Pardo (2019), Guvenen, Karahan, Ozkan, and Song (2021), Guvenen, Ozkan, and Madera (2024), and Kirkby (2025) for *labor earnings*.

[3] The nonlinear income tax in our paper applies only to labor income and captures tax credits and government transfers in the data as negative taxes. Although the actual income tax code also covers some forms of capital income, this type of income is often taxed separately from labor income, and various forms of capital income are subject to complex and heterogeneous tax rules. Therefore, following much of the literature on progressive income taxation we model the capital income tax as a separate and flat tax.

[4] When implementing the LIT, the government chooses a discount rate for calculating household lifetime earnings and tax liabilities; this rate is a policy parameter. We find that using a zero discount rate allows for larger welfare gains from a lifetime income tax than using the positive market interest rate. This choice has the additional advantage of removing age as argument from the tax formula.

There are two further unmodeled practical advantages of our implementation of a LIT. First, as long as the marginal tax rate is below 100%, households can always meet their tax obligations with their current-year earnings. This is satisfied with an AIT, but may not hold in alternative implementations of a LIT. Second, the government continues to collect taxes and settle accounts with each household on an annual basis, preventing the accumulation of tax liabilities over multiple years.

the same present discounted value of taxes from a newborn cohort.[5]

As in the simple model, both the annual and lifetime income taxes present policymakers with the trade-off between seeking to implement an efficient allocation of labor over time and across states on one hand, and the provision of social consumption insurance (against idiosyncratic risk) and redistribution across ex ante heterogeneous households on the other hand. With incomplete financial markets and in the absence of state-contingent lump-sum taxes, the efficiency and the insurance/redistribution motives cannot be fully separated in a competitive equilibrium. The desire to provide insurance and redistribution calls for a progressive income tax system. However, progressive income taxation entails rising marginal tax rates as earnings increase, which discourages labor supply in periods of high labor productivity and thus reduces aggregate labor efficiency, which we define as the ratio between aggregate earnings and aggregate hours worked. Furthermore, households can still achieve partial self-insurance through precautionary savings and adjustments in hours worked, and an expansion of public insurance may partially crowd out such private self-insurance.

The quantitative analysis confirms the theoretical prediction of the simple model that a LIT renders the labor supply distortions less severe, but at the expense of less effective social insurance and redistribution for a given degree of tax progressivity. Because marginal tax rates are less sensitive to earnings fluctuations under a lifetime income tax as long as the tax code is progressive, household labor supply decisions are less distorted over time and across idiosyncratic wage states, leading to a more efficient distribution of hours worked in the economy. Consequently, we find that switching from the status quo annual income tax to a lifetime income tax with the same progressivity raises aggregate labor efficiency by 1.23% and induces welfare gains equivalent to 0.63% of household lifetime consumption.[6]

Furthermore, we find that aggregate labor efficiency increases with tax progressivity (starting from the status quo) under a lifetime income tax, whereas it falls with an annual income tax. This is the result of two opposing forces. First, a more progressive tax system widens the gap in the marginal tax rate between high- and low-wage households, and thus depresses the incentives of high-wage households to work relative to low-wage households, in turn shifting the hours distribution away from high-productivity households, thereby

---

[5] That is, we deliberately abstract from intergenerational redistribution across different cohorts that a tax reform might entail in a stationary equilibrium. Insisting that all polices satisfy the same within-cohort government budget constraint insures that the net tax revenue collected from each cohort remains unchanged. If one further assumes that a policy reform is only applicable to new households and that factor prices are constant as we do in the benchmark analysis, then the transition (for newborn generations) is immediate and tax reforms showing welfare gains for newborn cohorts are in fact Pareto-improving: all current generations already alive are unaffected, and new generations immediately enjoy the welfare gains from the reform.

[6] We employ the two-parameter tax function from Bénabou (2002) and Heathcote, Storesletten, and Violante (2017) in which progressivity is measured as one minus the elasticity of after-tax earnings to pre-tax earnings.

reducing aggregate labor efficiency. This adverse effect is weaker, however, under a lifetime income tax since marginal tax rates are less responsive to current earnings under such a system. Second, however, a more progressive income tax system also reduces the dispersion of consumption (across ex-ante different households and across different states of the world ex post), and the relatively lower consumption of high-wage households induces them to work harder, thereby enhancing aggregate labor efficiency. Under a lifetime income tax this second effect dominates and aggregate labor efficiency actually rises with tax progressivity (up to a point), whereas it falls under annual income taxation.

The same intuition also carries over to our analysis of the optimal degree of tax progressivity. Due to the additional labor efficiency gains from a more progressive lifetime income tax, the optimal LIT is more progressive and attains larger welfare gains than the optimal AIT. Switching from the status quo policy to the optimal lifetime income tax boosts labor efficiency by 1.71% and improves social welfare by 1.27% of lifetime consumption. Consistent with our earlier discussion, the optimal LIT induces a more efficient distribution of hours worked, and therefore leads to greater hours- and earnings inequality. Nevertheless, this tax system reduces *consumption* inequality substantially. In comparison, the optimal AIT reduces labor efficiency by 0.19% and achieves a welfare gain equivalent to 0.19% of lifetime consumption relative to the current annual income tax with status quo progressivity.

Finally, accounting for general equilibrium effects from the tax reforms leads to significantly more progressive optimal policies and larger welfare gains for both annual and lifetime income taxes. However, the welfare ranking between the lifetime income tax and annual income tax remains unchanged in that the lifetime tax leads to maximum welfare gains that are about 1.55% (of lifetime consumption) higher than the annual tax (7.39% vs. 5.84%).

1.1 Related Literature

The idea of lifetime income taxes dates back to Vickrey (1939). In the context of progressive income taxation, and for the purpose of avoiding excessive taxes on fluctuating incomes relative to stable incomes and for preventing tax evasion via income shifting between years, Vickrey proposed an income tax system based on the average income of past years; our implementation of the LIT builds on this idea. Although Vickrey's proposal is close to a century old and has intuitive appeal, the literature on model-based quantitative analyses of progressive lifetime income taxation is sparse. Our paper seeks to partially fill this gap.

Our paper contributes to two broad literatures. First, a large body of work studies optimal nonlinear income tax in quantitative dynamic models with heterogeneous households in the Ramsey tradition in which a government can fully commit to a future path of taxes and is restricted to "simple" progressive tax functions. Examples include Bénabou (2002), Conesa and Krueger (2006), Conesa, Kitao, and Krueger (2009), Heathcote et al. (2017), HSV henceforth, and more recently Boar and Midrigan (2022), Dyrda and Pedroni (2022), and Holter, Krueger, and Stepanchuk (2023). This literature confines income tax policy to be based on annual income; our analysis uncovers that switching to a lifetime income tax holds the potential for significant welfare gains.

Second, the new dynamic public finance approach, developed as an alternative to the optimal Ramsey taxation literature, solves for the constrained efficient allocation in economies subject to informational and enforcement frictions, and then discusses the decentralization of these allocations with judiciously chosen tax systems, see, e.g., Golosov, Kocherlakota, and Tsyvinski (2003), Kocherlakota (2005), Albanesi and Sleet (2006), Battaglini and Coate (2008), Farhi and Werning (2007), Werning (2007), Farhi and Werning (2013), and Golosov, Troshkin, and Tsyvinski (2016). Typically, this decentralization requires a complex tax system that depends on the entire history of past incomes.[7] One actual part of the current U.S. fiscal constitution, as in many other countries, that features this type of history dependence is social security, although it differs from a lifetime income tax along several dimensions.[8] Consequently, our paper is related to the literature that studies the redistributive and incentive effects of social security. For example, Grochulski and Kocherlakota (2010) argue that it is possible to implement a socially optimal allocation using a social security system in which taxes/transfers are history-dependent only at retirement. Recent quantitative studies on the optimal progressivity of social security system with parameterized benefit function include Fehr, Kallweit, and Kindermann (2013) and Ábrahám, Brendler, and Cárceles-Poveda (2024). These papers examine a pay-as-you-go social security system, and intergenerational transfers play a crucial role in shaping their results. In contrast, we focus on comparing the performance of annual and lifetime income taxes in providing insurance and redistribution within each generation by imposing a within-cohort government budget constraint.[9]

The three papers most closely related to our study of a lifetime income tax are Huggett

---

[7] Starting from the observation that such history-dependent tax systems are complex to implement, a literature explores simplified alternatives retaining some key features of the full-history-dependent optimal tax policy, for example, by allowing taxes to depend on age, as studied by Kremer (2001), Erosa and Gervais (2002), Garriga (2001), Blomquist and Micheletto (2008), and more recently Heathcote, Storesletten, and Violante (2020). Weinzierl (2011) finds that introducing such age dependence into the income tax system can lead to substantial welfare gains. In contrast, our paper explores a different dimension of tax reforms and considers how large welfare gains can be achieved by introducing a simplified form of history dependence, i.e., allowing taxes to be conditioned on accumulated lifetime earnings.

[8] The actual U.S. system includes caps on taxable earnings, annuitization of benefits, and spousal and survivors benefits. Auerbach, Kotlikoff, and Koehler (2023) empirically measure inequality in lifetime spending power and the progressivity in the U.S. fiscal constitution including social security for shaping this inequality.

[9] Implementing a lifetime income tax through social security presents additional challenges. Since all taxes and transfers would be settled at retirement, households receiving transfers might need to borrow substantially early in life, and binding borrowing constraints may raise problems in this regard. Conversely, households with substantial tax liabilities might lack sufficient funds to pay these liabilities at retirement.

and Parra (2010), Kapička (2020), and Batzer (2022). The main focus of Huggett and Parra (2010) is a reform of the U.S. social security system. In the thought experiment most relevant for our paper, the authors replace, in a partial equilibrium analysis, the entire income tax and social security system with an optimal tax on lifetime earnings. We instead focus on a reform of the income tax system, keeping the social security system unchanged. They find that the introduction of a lifetime income tax leads to a small welfare loss, in the quantitative version of their model with persistent and transitory wage shocks. As Huggett and Parra (2010) acknowledge, however, their model understates wage and earnings inequality in the data.[10] This could lead to an understatement of the labor efficiency gains from a lifetime income tax that we stress in our paper. Our quantitative model aligns closely with the empirically observed life-cycle profiles of wage and earnings dispersion[11] and therefore has a larger scope for the lifetime income tax reform to generate labor efficiency- and welfare gains. Our quantitative results confirm that this potential indeed materializes.

Second, in his theoretical study of the optimal history dependence of income taxes, Kapička (2020) allows taxes to depend on a geometrically weighted average of past incomes. He then studies the optimal weight on past incomes, in the context of the analytically tractable incomplete markets economy of Heathcote, Storesletten, and Violante (2014). To obtain closed-form solutions, the paper abstracts from life cycle considerations and household savings decisions that are at the heart of our model. Furthermore, in his model optimal hours worked are constant over time and across labor productivity states. Therefore a change in tax policy only affects the level of labor supply, but cannot induce a more efficient distribution over the life cycle and across idiosyncratic states, the main source of welfare gains from a LIT in our model. In the absence of individual savings to smooth idiosyncratic shocks, a history-dependent income tax might be a useful tool to help households smooth consumption over time. Thus, Kapička (2020) finds that the optimal history-dependent tax is more progressive in current income but regressive in past incomes, i.e., a temporary increase in current income is taxed more heavily today but raises future after-tax income, in the same way private savings (if permitted) would respond to such shocks.[12]

Finally, the first chapter of Batzer (2022) studies optimal history-dependent income tax-

---

[10] They present inequality measures from their full model in Figure 4 of the original paper.

[11] The key difference between the two frameworks is that the estimated labor productivity in our model captures variation in wages due to *both* unobservable and observable factors; we also permit preference heterogeneity in the disutility of labor (as in Bick, Blandin, and Rogerson (2024) or Urquizo (2025)) which generates additional dispersion in hours worked.

[12] Further differences include the fact that in Kapička (2020) current income taxes depend on a weighted geometric average of current and past incomes, whereas our lifetime income tax is equivalent to using an arithmetic average. Finally, his paper permits the tax system to be age-dependent; the age-dependent tax level parameters are chosen optimally and are functions of the history-dependent tax parameters.

ation and considers a simplified implementation in which annual tax payments are restricted to follow an HSV tax function, with level and progressivity parameters that depend on both age and average past income. The main difference to our work is that Batzer (2022) emphasizes imperfect substitution across different types of labor in aggregate production and the resulting general equilibrium effects of income taxation on wages, whereas our paper focuses on the efficient allocation of labor over the life cycle and across idiosyncratic states. We view the two contributions as complementary, each highlighting important dimensions along which a LIT has advantages over a conventional AIT.

Section 2 presents an analytically tractable two-period model to explain the key benefits and costs of a progressive LIT, relative to a progressive AIT. Section 3 sets up the quantitative model, and Section 4 discusses its calibration. Section 5 investigates the welfare and economic implications of implementing a LIT. Section 6 conducts robustness analysis, and Section 7 concludes. The Appendix contains the proofs and further sensitivity analyses.

2 The Two-Period Model

Consider a two-period life-cycle model of households in partial equilibrium with exogenous interest rate and wages. There is a continuum of measure 1 of ex ante identical households that work and consume in each period, and they can also save or borrow at the risk-free interest rate r in the first period. In the first period, all households earn wage w_1, but the state in the second period $s \in S$ is uncertain, which affects the second-period wage $w_2(s)$. Let f and F denote the probability density function (pdf) and cumulative density function (cdf) for the atomless distribution of the state s. Shocks are idiosyncratic, and we assume that a law of large numbers applies, so that f and F are also the population distributions across households in the second period. Households value consumption and dislike labor according to the same period utility function we will assume in the quantitative model:

$$U(c,l) = \frac{c^{1-\sigma}-1}{1-\sigma} - \psi\frac{l^{1+1/\eta}}{1+1/\eta}. \tag{1}$$

The parameter σ governs household risk aversion, the parameter η is the Frisch elasticity of labor supply with respect to wage, and ψ controls the level of disutility from labor.

The government levies (potentially negative) taxes; for simplicity we assume that the government does not spend on goods. As private households, the government has access to an intertemporal technology that turns one unit of consumption today into one unit of consumption tomorrow. That is, the real interest in our model is equal to $r = 0$ both for private agents and the government (and the social planner). Furthermore, individuals (and thus the benevolent government) do not discount the future: the time discount rate is zero.

Let $j = 1, 2$ denote household age, c_j, l_j, and a denote household consumption, labor supply, and assets, respectively. The household's problem is then:

$$\max_{\{c_1,l_1,a,[c_2(s),l_2(s)]_{s\in S}\}} U(c_1,l_1)+\mathbf{E}\{U(c_2(s),l_2(s))\} \qquad \text{s.t.}$$

$$c_1+a=w_1l_1-T(w_1l_1), \quad \text{and} \quad c_2(s)=a+w_2(s)l_2(s)-\widetilde{T}(w_2(s)l_2(s),w_1l_1),\ \forall s\in S.$$

Expectations $\mathbf{E}$ are taken with respect to s, and T and $\widetilde{T}$ are tax functions in the first and second periods. Tax liabilities in the second period can depend on first-period earnings.

2.1 Tax Systems

Under an annual income tax (AIT), the tax function in the second period satisfies:

$$\widetilde{T}(w_2l_2,w_1l_1)=T(w_2l_2),$$

so the tax function is the same in each period and depends only on current-period earnings. The expected present value of taxes paid over the household lifetime is given by

$$T(w_1l_1)+\mathbf{E}\{T(w_2(s)l_2(s))\}.$$

In contrast, when the government levies a lifetime income tax (LIT), we assume that the second period tax bill is given by

$$\widetilde{T}(w_2l_2,w_1l_1)=T(w_1l_1+w_2l_2)-T(w_1l_1),$$

with expected present discounted value

$$T(w_1l_1)+\mathbf{E}\{\widetilde{T}(w_2(s)l_2(s),w_1l_1)\}=\mathbf{E}\{T(w_1l_1+w_2(s)l_2(s))\},$$

such that total lifetime tax liabilities are only a function of lifetime earnings $w_1l_1+w_2l_2$.

Throughout the paper we assume that the tax function $T(\cdot)$ is twice differentiable, and unless further noted, we assume that $T''(\cdot)>0$, implying that marginal tax rates are strictly increasing with income, and therefore the tax code is *progressive*. The parametric form of the tax function we will employ in our applied analysis, following Bénabou (2002) and Heathcote et al. (2017), is given by the two-parameter family:

$$T(y)=y-(1-\tau)y^{1-\mu}, \tag{2}$$

where μ and τ are two parameters governing the progressivity and the level of the income tax. Evidently T is twice differentiable, with $T''(y)>0$ for all $y>0$ as long as $\mu\in(0,1)$.

2.2 Efficient Allocation

As point of comparison to equilibrium allocations without and with taxes we first characterize the efficient allocation. This allocation is the solution to the social planner's problem where the planer chooses consumption and labor supply to maximize expected lifetime utility of the ex-ante identical households, subject to an economy-wide resource constraint:

$$\max_{\{c_1,l_1,[c_2(s),l_2(s)]_{s\in S}\}} U(c_1,l_1)+\mathbf{E}\{U(c_2(s),l_2(s))\} \qquad \text{s.t.}$$

$$c_1+\mathbf{E}\{c_2(s)\}=w_1l_1+\mathbf{E}\{w_2(s)l_2(s)\}. \tag{3}$$

The resource constraint (3) reflects the assumption that the planner also has access to the technology that can transfer resources between periods one for one. Note that the expectation **E** in the resource constraint is the expectation over productivity levels for each individual. With a law of large numbers this expectation is also the cross-sectional average across the continuum of households. The planner faces no aggregate risk and can freely allocate resources across individuals and thus across idiosyncratic states of the world s for any given individual. Proposition 1 provides a characterization of the efficient allocation.

Proposition 1 (Efficient Allocation)**.** *The efficient allocation* $\{c_1,l_1,[c_2(s),l_2(s)]_{s\in S}\}$ *is characterized by the following conditions:*

1. *Consumption Euler equation: for all* $s\in S$

$$U_c(c_1,l_1)=U_c(c_2(s),l_2(s)); \tag{4}$$

2. *Intratemporal optimality condition:*

$$-\frac{U_l(c_1,l_1)}{U_c(c_1,l_1)}=w_1, \quad \textit{and for all } s\in S, \quad -\frac{U_l(c_2(s),l_2(s))}{U_c(c_2(s),l_2(s))}=w_2(s), \tag{5}$$

and the resource constraint (3)*. The optimal allocation of labor is then characterized by:*

1. *Optimal allocation of labor between first and second period: for all* $s\in S$

$$\frac{U_l(c_1,l_1)}{w_1}=\frac{U_l(c_2(s),l_2(s))}{w_2(s)}; \tag{6}$$

2. *Optimal allocation of labor across states in the second period: for all* $s,s'\in S$

$$\frac{U_l(c_2(s'),l_2(s'))}{w_2(s')}=\frac{U_l(c_2(s),l_2(s))}{w_2(s)}. \tag{7}$$

If the period utility function is given by equation (1)*, the efficient allocation is given by*

$$c_1=c_2(s)=\left(\frac{\psi^{-\eta}}{2}\left[w_1^{1+\eta}+\mathbf{E}\left([w_2(s)]^{1+\eta}\right)\right]\right)^{\frac{1}{1+\sigma\eta}},$$

$$l_1=\left(\frac{\psi^{-\frac{1}{\sigma}}[w_1]^{\frac{1}{\sigma}+\eta}}{\frac{1}{2}\left[[w_1]^{1+\eta}+\mathbf{E}\left([w_2(s)]^{1+\eta}\right)\right]}\right)^{\frac{\sigma\eta}{1+\sigma\eta}} \quad \textit{and} \quad l_2(s')=\left(\frac{\psi^{-\frac{1}{\sigma}}[w_2(s')]^{\frac{1}{\sigma}+\eta}}{\frac{1}{2}\left[[w_1]^{1+\eta}+\mathbf{E}\left([w_2(s)]^{1+\eta}\right)\right]}\right)^{\frac{\sigma\eta}{1+\sigma\eta}}. \tag{8}$$

Proof. See Appendix A.1. □

Corollary 1. *If, in addition, $\sigma = 0$, then the efficient allocation of labor is given by*

$$l_1 = \left(\frac{w_1}{\psi}\right)^{\eta} \quad \text{and} \quad l_2(s') = \left(\frac{w_2(s')}{\psi}\right)^{\eta}, \quad \text{and thus}$$
$$\frac{l_2(s')}{l_1} = \left(\frac{w_2(s')}{w_1}\right)^{\eta} \quad \text{and} \quad \frac{l_2(s')}{l_2(s)} = \left(\frac{w_2(s')}{w_2(s)}\right)^{\eta} \quad \forall s, s'.$$

Proposition 1 states that the efficient labor allocation equates the marginal disutility of producing one unit of output U_l/w (which we refer to as the marginal disutility of earnings henceforth) across time and across states of the world. The social planner can separate this efficient allocation of labor from the efficient provision of consumption insurance, governed by the risk-sharing equations (4). Furthermore, if the utility function takes the form in (1), then labor supply is increasing in labor productivity in the current period and state, and decreasing in expected labor productivity over the life cycle (the denominator). In the absence of income effects on labor supply ($\sigma = 0$), the efficient allocation of labor is determined purely by current labor productivity: $l_1 = (w_1/\psi)^{\eta}$ and $l_2(s) = (w_2(s)/\psi)^{\eta}$ for all s.

2.3 The Laissez-Faire Equilibrium

We first consider laissez-faire equilibrium without government. Recall that the government does not need to raise taxes to pay for government consumption, and thus all taxes are zero in this section. Proposition 2 provides the conditions governing the equilibrium distribution of labor supply, which are counterparts of (6) and (7) for the efficient allocation.[13]

Proposition 2 (Laissez-Faire Equilibrium)**.** *The laissez-faire equilibrium allocation* $\{c_1, l_1, a, [c_2(s), l_2(s)]_{s\in S}\}$ *is characterized by*

1. *Consumption Euler equation*

$$U_c(c_1, l_1) = \mathbf{E}\{U_c(c_2(s), l_2(s))\}; \tag{9}$$

2. *Intratemporal optimality condition*

$$-\frac{U_l(c_1, l_1)}{U_c(c_1, l_1)} = w_1, \quad \text{and for all } s \in S, \quad -\frac{U_l(c_2(s), l_2(s))}{U_c(c_2(s), l_2(s))} = w_2(s), \tag{10}$$

and household budget constraints

$$c_1 + a = w_1 l_1, \quad \text{and for all } s \in S, \quad c_2(s) = w_2(s) l_2(s) + a.$$

---

[13] Note that in this partial equilibrium setting, the laissez-faire equilibrium is equivalent to the constrained efficient allocation, where the notion of constrained efficiency with incomplete asset markets is based on Dávila, Hong, Krusell, and Ríos-Rull (2012), i.e., it is the allocation chosen by a planner constrained from completing financial markets spanning idiosyncratic wage risk. This planner's problem imposes the additional constraint that consumption in the second period must be implementable with state non-contingent saving in the first period. In general equilibrium, the equivalence between competitive equilibrium- and constrained-efficient allocations no longer holds because the planner internalizes the effect of labor and savings allocations on the marginal product of labor and capital whereas households in competitive equilibrium do not.

These conditions imply the following equilibrium conditions for the allocation of labor:

1. *Allocation of labor between first and second period:*

$$\frac{U_l(c_1,l_1)}{w_1} = \mathbf{E}\left\{\frac{U_l(c_2(s),l_2(s))}{w_2(s)}\right\}; \tag{11}$$

2. *Allocation of labor across states in the second period: for all $s, s' \in S$,*

$$\frac{U_l(c_2(s'),l_2(s'))/w_2(s')}{U_l(c_2(s),l_2(s))/w_2(s)} = \frac{U_c(c_2(s'),l_2(s'))}{U_c(c_2(s),l_2(s))}. \tag{12}$$

Proof. See Appendix A.2. ☐

In comparison to the efficient allocation, Proposition 2 states that in the laissez-faire equilibrium, the marginal disutility of earnings U_l/w is equated only in expectation, as stipulated in equation (11), rather than state-by-state as in (6). This reflects the missing insurance markets against the idiosyncratic risk, and hence households cannot adjust labor supply without changing their consumption in the same contingent way. Consequently, even if the marginal disutility of earnings is higher in one state, households might not shift labor supply away from this state because the associated consumption in that state would fall too much. This is evident from equation (12) which links the relative marginal disutility of earnings to the marginal utility of consumption in the laissez-faire equilibrium. In particular, the higher is the marginal utility of consumption (i.e., the lower is consumption) in a given state, the more households are willing to tolerate a higher marginal disutility of earnings (i.e., the more they work). The marginal disutility of earnings is still equalized over time in expectation because households can shift resources over time through state-non-contingent saving.

Note that when households are risk-neutral ($\sigma = 0$), this interaction between labor supply and desired self-insurance of consumption is absent, as the following corollary shows.

Corollary 2. *If the utility function is of the form* (1) *and $\sigma = 0$, then the laissez-faire equilibrium allocation of labor $\{l_1, [l_2(s)]_{s\in S}\}$ is given in Corollary 1 and thus is efficient.*

In the presence of missing insurance markets, however, by taxing households based on their earnings, a government may improve on the laissez-faire equilibrium by providing partial insurance against the idiosyncratic risk, but at the cost of further distorting the distribution of labor supply relative to the efficient benchmark in Proposition 1. That is, the government might want to tolerate additional distortions on labor supply in exchange for better insurance. This trade-off is the same in the presence of period taxation and lifetime income taxation, but how strong the two effects are, and what they imply for the optimal progressivity of the income tax code differ across the two tax systems.

2.4 Annual vs. Lifetime Income Taxes: Theory

We now characterize the equilibrium labor allocations under an AIT and a LIT in the two-period model and compare them to the efficient allocation in Section 2.2 and the laissez-faire equilibrium in Section 2.3. Proposition 3 contains the results which are the counterparts of Proposition 1 for the efficient allocation and Proposition 2 for the laissez-faire equilibrium.

Proposition 3 (Annual vs. Lifetime Income Taxes)**.** *For a given tax system, the equilibrium allocation is characterized by the distribution of labor supply:*

1. *Allocation of labor between first and second period:*
   (a) *Annual income tax:*
$$\frac{U_l(c_1, l_1)}{w_1[1 - T'(w_1 l_1)]} = \mathbf{E}\left\{\frac{U_l(c_2(s), l_2(s))}{w_2(s)[1 - T'(w_2(s) l_2(s))]}\right\}; \tag{13}$$
   (b) *Lifetime income tax:*
$$\frac{U_l(c_1, l_1)}{w_1} = \mathbf{E}\left\{\frac{U_l(c_2(s), l_2(s))}{w_2(s)}\right\}. \tag{14}$$
2. *Allocation of labor across states in the second period: $\forall s, s'$,*
   (a) *Annual income tax:*
$$\frac{U_l(c_2(s'), l_2(s'))/w_2(s')}{U_l(c_2(s), l_2(s))/w_2(s)} = \left[\frac{1 - T'(w_2(s') l_2(s'))}{1 - T'(w_2(s) l_2(s))}\right] \frac{U_c(c_2(s'), l_2(s'))}{U_c(c_2(s), l_2(s))}; \tag{15}$$
   (b) *Lifetime income tax:*
$$\frac{U_l(c_2(s'), l_2(s'))/w_2(s')}{U_l(c_2(s), l_2(s))/w_2(s)} = \left[\frac{1 - T'(w_1 l_1 + w_2(s') l_2(s'))}{1 - T'(w_1 l_1 + w_2(s) l_2(s))}\right] \frac{U_c(c_2(s'), l_2(s'))}{U_c(c_2(s), l_2(s))}. \tag{16}$$

Proof. See Appendix A.3. □

First, under both annual and lifetime income taxes, the marginal disutility of earnings is equated across time only in expectation, as opposed to holding state-by-state in the efficient allocation in equation (6). This is again the consequence of incomplete insurance markets against idiosyncratic wage risk and the assumed absence of state-contingent lump-sum transfers/taxes. Furthermore, equation (13) shows that with an AIT, it is the marginal disutility of *after-tax* earnings that in expectation is equated over time, instead of the marginal disutility of (pre-tax) earnings in the laissez-faire equilibrium. Consequently, variations in the marginal tax rate across periods distort labor supply intertemporally. If the income tax is progressive and thus the marginal tax rate is increasing in labor income, then labor supply in high-productivity (and thus high income) periods is depressed relative to that in low-productivity periods under the AIT. In contrast, (14) shows that a LIT, even when progressive, introduces no distortion along this margin and labor supply satisfies the same intertemporal optimality condition as in the laissez-faire equilibrium, see (11). In the absence of income risk, this observation also unambiguously implies that aggregate labor efficiency, measured as total earnings divided by total hours, is higher under this tax system, and this is the key advantage of a LIT. The following corollary summarizes this discussion.

Corollary 3. *Suppose there is no income risk in the second period, and $w_1 \neq w_2$. Then the intertemporal allocation of labor is efficient under the LIT and aggregate labor efficiency $ALE = \frac{w_1 l_1 + w_2 l_2}{l_1 + l_2}$ is strictly higher under the LIT than under the AIT.*

With uninsurable income risk in the second period both an AIT and LIT add additional distortions to household labor supply, relative to the laissez-faire equilibrium. Comparing equations (15) and (16) to (12), with distortionary taxes the marginal disutility of earnings across states is not only linked to the marginal utility of consumption (as in laissez-faire), but also to the state-contingent marginal tax rate. Both income taxes reduce relative labor supply when marginal tax rates are high, that is (with a progressive tax code), when earnings are high. The difference in marginal tax rate across states tends to be smaller under a LIT because the marginal tax rate is determined by lifetime earnings rather than current earnings, and these lifetime earnings are less sensitive to the state-contingent earnings in the current period. The distribution of labor supply across contingent states then follows that of labor productivity (as the efficient allocation would stipulate) more strongly when the tax base is lifetime income than when it is annual income. Thus, the simple model suggests that a LIT tends to induce a more efficient distribution of labor supply than an AIT when taxes are progressive: it introduces no distortion intertemporally and less distortions across states. Households will work longer hours when their labor productivity is high under a LIT than under an AIT, and aggregate labor efficiency is higher under this tax system.

If individuals are risk-neutral ($\sigma = 0$) and there are no income effects on labor supply, the response of labor to labor productivity is unambiguously closer to the efficient (and laissez-faire) benchmark under the LIT than under the AIT, as the following corollary states.

Corollary 4. *If the utility function is of the form (1), the tax function of the form (2) and $\sigma = 0$, then the equilibrium labor allocation $\{l_1, [l_2(s)]_{s \in S}\}$ under the AIT and the LIT satisfy:*

(a) Annual income tax: for all s, s'

$$\frac{l_2(s')}{l_1} = \left(\frac{w_2(s')}{w_1}\right)^{\eta\left(\frac{1-\mu}{1+\eta\mu}\right)} \qquad \text{and} \qquad \frac{l_2(s')}{l_2(s)} = \left(\frac{w_2(s')}{w_2(s)}\right)^{\eta\left(\frac{1-\mu}{1+\eta\mu}\right)};$$

(b) Lifetime income tax: for all s, s'

$$\frac{l_2(s')}{l_1} = \left(\frac{w_2(s')}{w_1}\right)^{\eta} \frac{\left(w_1 l_1 + l_2(s') w_2(s')\right)^{-\mu\eta}}{\left[\mathbf{E}\left(w_1 l_1 + l_2(\tilde{s}) w_2(\tilde{s})\right)^{-\mu}\right]^{\eta}} \quad \text{and} \quad \frac{l_2(s')}{l_2(s)} = \left(\frac{w_2(s')}{w_2(s)}\right)^{\eta} \left(\frac{w_1 l_1 + l_2(s') w_2(s')}{w_1 l_1 + l_2(s) w_2(s)}\right)^{-\mu\eta}.$$

Elasticities of labor to productivity $\varepsilon_{2,1} = \frac{d\log(l_2(s')/l_1)}{d\log(w_2(s')/w_1))}$ *and* $\varepsilon_{s',s} = \frac{d\log(l_2(s')/l_2(s))}{d\log(w_2(s')/w_2(s)))}$ *satisfy*[14]

$$\eta = \varepsilon^{E}_{2,1} = \varepsilon^{LF}_{2,1} \geq \varepsilon^{LIT}_{2,1} > \varepsilon^{AIT}_{2,1} = \eta\left(\frac{1-\mu}{1+\eta\mu}\right) \quad \textit{and} \quad \eta = \varepsilon^{E}_{s',s} = \varepsilon^{LF}_{s',s} \geq \varepsilon^{LIT}_{s',s} > \varepsilon^{AIT}_{s',s} = \eta\left(\frac{1-\mu}{1+\eta\mu}\right),$$

where superscripts E and LF denote efficient and laissez-faire, respectively.

However, the property that a LIT distorts labor supply less (or not at all) across periods can be a disadvantage when households are risk averse ($\sigma > 0$) and market incompleteness prevents explicit consumption insurance; in contrast to the planner solution, the efficient allocation of labor cannot be separated from the efficient provision of consumption insurance. The lower is second-period labor supply, the less relevant are idiosyncratic uninsurable wage shocks for consumption risk. Therefore, for consumption insurance purposes it could be welfare improving to impose higher tax rates during high-earnings periods. This is, however, impossible under a LIT because, as equation (14) shows, this tax system does not distort labor supply intertemporally. There is then a nontrivial trade-off between the labor efficiency benefits and the consumption insurance costs of the LIT relative to the AIT.

In the remainder of this paper we quantify this trade-off. In Appendix B we provide a quantitative analysis of the simple model, in order to identify the main quantitative determinants of this trade-off. There we show that although the allocation of labor supply is closer to the efficient allocation under a LIT (as predicted by Proposition 3 and Corollary 4), the AIT can mimic an age-dependent labor income tax, and tax the stages of the life cycle with greater risk more heavily, therefore providing better consumption insurance. In addition, even though the AIT distorts the allocation of labor supply more across states of the world, as a by-product it reduces earnings inequality and in turn lowers consumption inequality. See Figure 8 in the appendix for this trade-off. There, we also show (see Figure 9 in Appendix B) that when the progressivity of the labor income tax code increases, the labor efficiency advantage of the LIT rises. This in turn is the main reason why the optimal LIT is more progressive than the optimal AIT, a pattern in the simple model that will also be a central feature of the quantitative life-cycle model with intertemporal labor supply, precautionary saving, and social insurance. We turn to this analysis in the main text next.

3 The Quantitative Model

We now introduce our quantitative life-cycle model with heterogeneous households through which we interpret the data and evaluate the economic and welfare consequences of tax re-

---

[14] We compute these elasticities by varying $w_2(s')$, holding w_1 and $w_2(s)$ constant. Since the distribution of s was assumed to be atomless, a change in $w_2(s')$ and associated change in $l_2(s')$ has no impact on $\mathbf{E}(\cdot)$. In the absence of risk, the expectation is trivial, $\varepsilon^{LIT}_{2,1} = \eta$ and the intertemporal allocation of labor is efficient. In this case the results of Corollary 3 (and Corollaries 1 and 2) extend fully to a model with $N > 2$ periods.

form. We first describe the maximization problems of households and firms, the role of government and two types of income tax policies, and then define a competitive equilibrium.

3.1 Households

Consider a stationary economy populated by overlapping generations of heterogeneous households. In each period, a continuum of measure one households is born at age 1. Households work in the first J_R years of their life cycles, then retire and live up to a maximum age of J. They face age-specific mortality risk; the survival probability for an age-j household is denoted by γ_{j+1}. We now describe the household optimization problems in recursive formulation, indexing the current household state by $\mathbf{S}$ and next period's state by $\mathbf{S}'$.

3.1.1 Working Households

Between age 1 and age J_R, households can work and earn labor income y determined by their labor supply l and wage rate $\widetilde{w}$. The household's wage $\widetilde{w}$ is given by

$$\widetilde{w}(j, z, \varepsilon; w) = w \exp\{\widetilde{e}(j) + z + \varepsilon\},$$

where w is the price per effective unit of labor, $\widetilde{e}(j)$ is a deterministic life-cycle trend of labor productivity that is common among households, and (z, ε) are the persistent and transitory idiosyncratic components of labor productivity. Following Arellano et al. (2017)'s study of *earnings* dynamics, we use a stochastic model based on quantile functions for idiosyncratic *wage* risk. It can capture the non-linear, non-Gaussian features of the data emphasized in the recent literature on *earnings* dynamics, see De Nardi et al. (2019) (who use the same structure as we do), Guvenen et al. (2021), Guvenen et al. (2024), and Kirkby (2025). The stochastic process for z is given by

$$z_j = Q_z(z_{j-1}, j, u_{z_j}), \quad u_{z_j} \sim \text{i.i.d. Unif}(0,1),$$

where $Q_z(\cdot)$ is the quantile function for the distribution of z_j conditional on z_{j-1} and age j, and u_{z_j} is an i.i.d. uniformly distributed draw that determines the rank of z_j in this conditional distribution.[15] For example, if $u_{z_j} = 0.5$, then z_j is at the median of the (conditional on z_{j-1} and j) distribution. The initial age 1 distribution of z is given by

$$z_1 = Q_{z_1}(j = 1, u_{z_1}), \quad u_{z_1} \sim \text{i.i.d. Unif}(0,1),$$

where $Q_{z_1}(j = 1, \cdot)$ is the quantile function for the distribution of z at age $j = 1$.

---

[15] A conventional AR(1) process is a special case of this more general formulation, with quantile function:

$$Q_z(z_{j-1}, j, u_{z_j}) = \rho z_{j-1} + \Phi^{-1}_{(0,\sigma^2_\nu)}(u_{z_j}),$$

where ρ is the persistence parameter, σ^2_ν is the variance of the persistent shock, and $\Phi^{-1}_{(0,\sigma^2_\nu)}(\cdot)$ is the inverse CDF of a normal distribution with zero mean and variance σ^2_ν. This persistent component z corresponds to the persistent or permanent (or a combination thereof) wage component in the extant literature.

The transitory component ε is independent across households and over time with zero mean. Its distribution is given by

$$\varepsilon = Q_\varepsilon(j, u_\varepsilon), \quad u_\varepsilon \sim \text{i.i.d. Unif}(0,1),$$

where $Q_\varepsilon(\cdot)$ is the quantile function for the distribution of ε at age j, and u_ε again is an i.i.d. uniform draw for its rank.

Financial markets are incomplete; households can only borrow and save in a risk-free asset at interest rate r, subject to age-dependent and potentially binding borrowing constraints $\underline{a}_{j+1}$. Households draw their initial assets from a non-degenerate distribution whose form we specify in the calibration section 4. Insurance contracts that pay out contingent on idiosyncratic labor productivity realizations are absent by assumption, as standard in the incomplete markets literature, see Bewley (1986), Huggett (1993), and Aiyagari (1994).

Four types of taxes are imposed by the government on working households: a labor income tax, a payroll tax, a capital income tax, and a consumption tax. The labor income tax policy is summarized by a function $\widetilde{T}(y,Y)$, which gives the current tax liability based on household current earnings y and, in the case of a LIT, the accumulated lifetime earnings Y *prior to the current period*. Payroll taxes, capital income taxes, and consumption taxes $(\tau_{ss}, \tau_k, \tau_c)$ are modeled as flat taxes on earnings, capital income, and consumption, respectively.

Household preferences are represented by the additively separable (between consumption and labor) period utility function given by (1). Following Kaplan (2012) and Heathcote et al. (2017), we permit heterogeneity in the disutility of labor, ψ, to match earnings and hours inequality in the data. Thus, we denote period utility by $u(c,l;\psi)$.

In each year, working households choose consumption c, labor supply l, and savings a' based on their current state $\mathbf{S} = \{a, Y, z, \varepsilon, \psi, j\}$, where a denotes the household's current assets. A working household's decision problem in the recursive formulation is then:

$$V(\mathbf{S}) = \max_{\{c,l,a',Y'\}} \left\{U(c,l;\psi) + \beta\gamma_{j+1}\mathbf{E}_{(z',\varepsilon')}\left[V(\mathbf{S}')|z,j\right]\right\} \qquad \text{s.t.}$$

$$(1+\tau_c)c + a' = y - \widetilde{T}(y,Y) - \tau_{ss}y + [1 + (1-\tau_k)r]a;$$

$$y = [w \exp\{\widetilde{e}(j) + z + \varepsilon\}]\, l; \;\; Y' = Y + y;$$

$$z' = Q_z(z, j+1, u_z), \;\; u_z \sim \text{Unif}(0,1); \;\; \varepsilon' = Q_\varepsilon(j+1, u_\varepsilon), \;\; u_\varepsilon \sim \text{Unif}(0,1); \;\; a' \geq \underline{a}_{j+1}, \;\; c \geq 0, \;\; l \geq 0,$$

where β is the time discount factor.[16]

---

[16] In our baseline analysis, the discount rate used to compute accumulated lifetime earnings is set to zero, so that $Y' = Y + y$ and the lifetime income tax function $\widetilde{T}(y,Y)$ is age-independent. Section 6.1 explores the implications of alternative discount rate choices used by the government.

3.1.2 Retired Households

Households retire at age J_R and receive retirement benefits b from the government in each year that depend on their average earnings $\bar{y}$ across working years. Benefits are determined by the benefit function $b = \widetilde{b}(\bar{y})$ where average earnings $\bar{y}$ are implicitly defined by:

$$Y_{J_R} = \sum_{j=1}^{J_R} y_j = \sum_{j=1}^{J_R} \bar{y}. \tag{17}$$

Retired households consume and save in the risk-free bond, and pay consumption- and capital income taxes. Since retirement benefits are fixed once determined, the state space of retired households is only $\mathbf{S} = \{a, b, j\}$ and the associated decision problem is

$$\begin{aligned} V^R(\mathbf{S}) \;&=\; \max_{\{c,a'\}} \left\{ U(c,0) + \beta\gamma_{j+1} V^R(a', b, j+1) \right\} \qquad \text{s.t.} \\ (1+\tau_c)c + a' \;&=\; b + [1 + (1-\tau_k)r]a;\;\; a' \geq \underline{a}_{j+1},\;\; c \geq 0. \end{aligned}$$

3.2 Firms

The production side of the economy consists of profit-maximizing firms that rent capital K at rate $r + \delta$ (where δ is the capital depreciation rate) and hire effective labor N at price w to produce the final output good. All firms produce according to a constant-returns-to-scale production technology $F(K, N)$, and all markets are perfectly competitive. Thus, without loss of generality we can focus on a representative firm that takes prices as given and maximizes per-period profits (which will be zero in equilibrium):

$$\max_{(K,N)} F(K, N) - (r + \delta)K - wN.$$

3.3 The Government

The government collects labor income taxes, payroll taxes, capital income taxes, consumption taxes, and accidental bequests from households to finance four types of expenditures: (1) retirement benefits, (2) expenditures on public goods, (3) interest payments on government debt, and (4) initial net worth of newborn households.

3.3.1 Annual vs. Lifetime Income Tax

We consider two types of labor income tax systems. Firstly, an annual income tax (AIT) as currently adopted by most countries. Under such policy, annual tax liabilities of each household depend only on current-year earnings y, and the tax function reduces to

$$\widetilde{T}(y, Y) = T(y),$$

where $T(\cdot)$ specifies the mapping from a household's current earnings y to its current tax liability $T(y)$. The second type of income tax system is a lifetime income tax (LIT) in which the total tax liability of a household depends on accumulated earnings over the life cycle. There are different ways of implementing such a LIT; a simple one is to set the tax function

$$\widetilde{T}(y, Y) = T\left(Y + y\right) - T(Y),$$

where $T(\cdot)$ now specifies the mapping from a household's lifetime earnings to their total tax liability. That is, at age j, households only pay the increment in their total tax liability due to the addition of their current-year earnings y to their accumulated lifetime earnings Y prior to the current period. It is easy to verify that the total value of all tax payments over a household's working life is given by

$$\sum_{j=1}^{J_R} \widetilde{T}(y_j, Y) = T\left(\sum_{j=1}^{J_R} y_j\right),$$

and is a function only of total value of realized labor earnings over the working life.[17]

3.3.2 Government Budget Constraint

The stationary government period budget constraint across different age cohorts is:

$$\sum_{j=1}^{J} \int \left[\widetilde{T}(y(\mathbf{S}), Y) + \tau_{ss} y(\mathbf{S}) + \tau_c c(\mathbf{S}) + \tau_k r a - b\right] d\Phi_j(\mathbf{S}) \\ + (1+r)\left[\sum_{j=1}^{J} \int (1-\gamma_{j+1}) a'(\mathbf{S}) d\Phi_j(\mathbf{S}) - \int a d\Phi_1(\mathbf{S})\right] = G + rB, \quad (18)$$

where $\Phi_j(\mathbf{S})$ is the measure of age-j households with state $\mathbf{S}$. G denotes government expenditures on public goods, and B is the amount of government debt. Household earnings y, lifetime earnings Y, consumption c, future assets a', and retirement benefit b are all functions of the household state $\mathbf{S}$. The left-hand side of the budget constraint is the government's net revenues from households, i.e., taxes minus retirement benefits plus accidental bequests minus the initial net worth of newborn households,[18] and the right-hand side includes government expenditures on public goods and interest payments on government debt.

3.4 Definition of Stationary Competitive Equilibrium

Given government tax policy $\{\widetilde{T}(\cdot), \tau_{ss}, \tau_k, \tau_c\}$, retirement benefit function $\widetilde{b}(\cdot)$, government debt B, and the measure of states for newborn households $\Phi_1(\cdot)$, a stationary competitive equilibrium is a collection of household value and policy functions $\{V, c, l, a'\}$, the representative firm's decisions $\{K, N\}$, expenditures on public goods G, the price of effective labor

---

[17] This implementation of the lifetime income tax has two desirable properties: first, as long as the average tax rate is below 100%, households can always afford to pay taxes with their current earnings. Second, the government still collects taxes at annual frequency, and no household accumulates tax liabilities over time.

[18] We assume that accidental bequests fund the initial assets of newborn households, with any remaining bequests accruing to the government as revenues, akin to estate taxes. The government earns interest on accidental bequests, which explains the factor $1 + r$ in the budget constraint (18).

w, interest rate r, and a sequence of measures for the household state $\{\Phi_j(\cdot)\}_{j=2}^{J}$ such that

1. Given prices $\{w, r\}$, tax policy $\{\widetilde{T}(\cdot), \tau_{ss}, \tau_k, \tau_c\}$ and retirement benefit function $\widetilde{b}(\cdot)$, the value and policy functions $\{V, c, l, a'\}$ solve the household optimization problem.
2. Given $\{w, r\}$, choices $\{K, L\}$ solve the representative firm's maximization problem.
3. Given the retirement benefit function $\widetilde{b}(\cdot)$, government debt B and government spending G, and given household policy functions $\{c, l, a'\}$ and population measures $\{\Phi_j(\cdot)\}_{j=1}^{J}$, the tax policy $\{\widetilde{T}(\cdot), \tau_{ss}, \tau_k, \tau_c\}$ satisfies the government period budget constraint (18).[19]
4. The labor market, capital market, and goods markets clear:

$$N = \sum_{j=1}^{J} \int \exp\{\widetilde{e}(j) + z + \varepsilon\} l(\mathbf{S}) d\Phi_j(\mathbf{S}); \qquad K + B = \sum_{j=1}^{J} \int a d\Phi_j(\mathbf{S});$$

$$\sum_{j=1}^{J} \int c(\mathbf{S}) d\Phi_j(\mathbf{S}) + G = F(K, N) - \delta K.$$

5. Given $\Phi_1(\cdot)$, the laws of motion for $\{\Phi_j(\cdot)\}_{j=1}^{J}$ induced by the household policy functions, demographics, and idiosyncratic shocks, $\{M_j(\cdot)\}_{j=1}^{J-1}$, satisfy $\Phi_{j+1} = M_j(\Phi_j)$ for all j.

4 Calibration

In this section, we describe how we parameterize the model to map it into U.S. data. We first introduce the main data sources, and then explain our calibration strategy and report the calibrated values of parameters. Lastly, we evaluate the model's performance in replicating the empirical life cycles of household variables and their dispersions.

4.1 Data

Our main data source is the core sample of the 1999-2017 PSID, from which we obtain individual and household level information about earnings, hours worked, consumption, net worth, and characteristics such as age, race, education, number of children. Since our model does not differentiate between single and married households, all household variables are normalized by the number of adults in the household (head and spouse only) before comparing to their model counterparts. Wages are constructed as earnings divided by hours worked.[20]

---

[19] We use the government period budget constraint (18) to calibrate public spending G, and employ it in our analyses of transitional dynamics, general equilibrium, and robustness. In the baseline *normative* policy exercises, however, we impose a *within-cohort* government budget constraint, requiring all tax policies to raise the same revenue from a newborn cohort. In general, we use the across-cohort period budget constraint (18) for positive analysis, viewing the data as generated by a world with interacting cohorts. For normative analysis, however, this constraint can imply large intergenerational redistribution towards newborns (see Appendix C.7). For the welfare analysis we therefore adopt the within-cohort constraint, which neutralizes such redistribution and provides a conservative benchmark for comparing welfare between the LIT and AIT.

[20] All nominal variables are converted to values in 2016 U.S. dollars based on the consumer price index for all urban consumers (CPI-U) from the U.S. Bureau of Labor Statistics. One unit of income in the model corresponds to the average earnings of households aged 25-60 in the PSID sample, which is \$54,248 per adult.

4.2 Calibration Strategy

The stationary competitive equilibrium in our model is calibrated to mirror the U.S. economy. Model moments are computed from the approximated household distribution implied by household policy functions and laws of motion for the exogenous variables.

4.2.1 Demographics

A model period represents one year in the data. Age 1 corresponds to data age 25, and the maximum age J is set to 76, i.e., data age 100. Age-specific survival probabilities γ_j are computed from the mortality rates published by the SSA. Figure 10 of Appendix C presents the life-cycle profile of survival probabilities and population shares. The implied life expectancy is 78.99. The retirement age J_R is set to 42 in the model, i.e., data age 66.

4.2.2 Preferences

For the distribution of labor disutility ψ, we assume $\ln \psi \sim N\left(\ln \bar{\psi}, \sigma_\psi^2\right)$ in the population, where $\bar{\psi}$ controls the level of disutility and σ_ψ^2 its dispersion. The value of $\bar{\psi}$ is calibrated to match average earnings of age 25-60 households, which is normalized to one in the model economy. σ_ψ^2 is pinned down by the covariance between log earnings and log hours worked in the PSID sample, which is 0.096.[21]

The value of σ in the literature typically ranges between 1 (i.e., log utility) and 2, and thus we choose the middle value of 1.5. We set the labor supply elasticity η to 0.5, broadly consistent with the microeconomic evidence on the Frisch elasticity. The discount factor β is calibrated to match the average net worth of age 51-60 households, which is 5.405 in model units, or $293,194 per adult in 2016 dollars, according to the PSID sample.

4.2.3 Wage Process

For the deterministic life-cycle trend of labor productivity $\widetilde{e}(j)$, we regress log-wages from the 1999-2017 PSID on age and year dummies and approximate the predicted log-wage at each age with a third degree polynomial. Figure 11 of Appendix C presents the results. For the idiosyncratic persistent and transitory wage components, we estimate three quantile functions: $Q_z(\cdot)$ for the conditional distribution of the persistent shock given its previous value and age, $Q_{z_1}(\cdot)$ for the initial distribution of the persistent component, and $Q_\varepsilon(\cdot)$ for the distribution of the transitory shock. The estimation follows Arellano et al. (2017) (ABB).

The quantile functions are parameterized using low-order Hermite polynomials:

$$Q_z(z_j, j, u_{z_j}) = \sum_{m=1}^{M} a_m^{Q_z}(u_{z_j}) \varphi_m^{Q_z}(z_j, j), \quad Q_x(j, u_x) = \sum_{m=1}^{M} a_m^{Q_x}(u_x) \varphi_m^{Q_x}(j), \;\; x \in \{z_1, \varepsilon\},$$

where $\varphi_m^{Q_x}(\cdot)$ are products of polynomials, and $a_m^{Q_x}(\cdot)$ are coefficients. The coefficients are specified as piecewise-linear functions of the shock percentiles (i.e., u_{z_j}, u_{z_1}, u_ε), augmented

---

[21] We target this covariance since it is immune to classical measurement errors in earnings and hours.

with parametric tails based on exponential distributions for the constant terms. Following ABB, we use tensor products of Hermite polynomials of degrees (3,2) in the previous value and age for $Q_z(\cdot)$, second-order polynomials of age for $Q_{z_1}(\cdot)$ and $Q_\varepsilon(\cdot)$, and 11 knots of percentiles for the piecewise-linear coefficient functions.[22]

Given our goal of comparing the performance of annual versus lifetime income taxes, it is essential that the wage process is estimated at the *annual* frequency. Therefore, we use the 1968-1993 PSID annual data (instead of the 1999-2017 biennial data), as De Nardi et al. (2019) did for an income process.[23] Figures 12-15 in Appendix C show that our estimated wage process fits the data well, and is consistent with the non-linear and non-Gaussian features emphasized in recent literature on income dynamics (Arellano et al. 2017; De Nardi et al. 2019; Guvenen et al. 2021, 2024; Kirkby 2025). Notably, wage persistence depends on age and the percentile ranks of previous wage and shock received, and the distribution of wage growth exhibits excess kurtosis relative to a normal distribution.

To solve the life-cycle model with the estimated process, we follow the discretization method in De Nardi et al. (2019) and construct Markov chains with age-dependent grids and transition matrices[24] for the persistent and transitory components.[25]

4.2.4 Government Policies

Income, capital, and consumption taxes. For the status quo economy with an AIT we approximate the U.S. tax-and-transfer system with the tax function in equation (2). The values of the level and progressivity parameters (τ, μ) are set based on the estimates by Wu (2021), which combine the federal and state income taxes and include government transfers as negative taxes. The tax progressivity parameter μ is 0.137 and the level parameter τ is set to 0.105 such that the average income tax rate of the median income household is 7.8%.

---

[22] This parameterization results in $(12 \times 11 + 2) + (3 \times 11 + 2) \times 2 = 204$ parameters, which are estimated using a stochastic EM algorithm augmented with quantile regressions. Details of the estimation procedure we follow are contained in ABB.

[23] We use the 1968-1993 PSID because the survey became biennial starting in 1999, and definitions of income variables changed after 1993. The wage process is estimated using residual log-wages after removing age and time effects from log-wage levels. The sample is restricted to household heads aged 25 to 60 with annual hours between 520 and 5200 and wages above one-half of the minimum wage.

[24] To capture the higher wage inequality in the 2000s, we scale up the grids for the wage process estimated from the 1968–1993 PSID annual data to match the life-cycle profile of the variance of log-wage in the 1999-2017 PSID biennial data. Figures 16-19 in Appendix C show that the discretized wage process captures the features of the data reasonably well, with the scaling up of the grids having almost negligible effects, except for the dispersion of wage growth. Figure 20 shows that the model also matches the *two-year* wage growth distributions in the 1999-2017 PSID biennial data, despite not targeting them in the estimation.

[25] Following the literature, the variance of measurement error in wage is set to 0.02 when comparing the model to data. Since wages are computed as earnings divided by hours worked, the variance of measurement error in wages follows from our assumption of zero measurement error in earnings and classical measurement error with variance 0.02 in log hours. The assumption of zero measurement error in earnings is consistent with the findings in Kaplan (2012) and Heathcote et al. (2014).

The capital income tax rate τ_k and consumption tax rate τ_c are also from Wu (2021), estimated using OECD aggregate data, following closely the method in Mendoza, Razin, and Tesar (1994) and Trabandt and Uhlig (2011). The capital income tax rate τ_k is 33.0%, and the consumption tax rate τ_c is 4.1%.

Payroll tax and retirement benefit. The payroll tax rate τ_{ss} is set to 12.4% based on the actual Social Security contribution rate. Like the U.S. social security system, the retirement benefit in the model is a piecewise-linear function of the household's average earnings before retirement $\bar{y}$, defined in equation (17):

$$\widetilde{b}(\bar{y}) = \begin{cases} \xi_{b,1}\bar{y} & \text{if } \bar{y} \leq b_1; \\ \xi_{b,1}b_1 + \xi_{b,2}(\bar{y} - b_1) & \text{if } b_1 < \bar{y} \leq b_2; \\ \xi_{b,1}b_1 + \xi_{b,2}(b_2 - b_1) + \xi_{b,3}(\bar{y} - b_2) & \text{if } b_2 < \bar{y} \leq b_3; \\ \xi_{b,1}b_1 + \xi_{b,2}(b_2 - b_1) + \xi_{b,3}(b_3 - b_2) & \text{if } \bar{y} > b_3. \end{cases}$$

Following the actual U.S. policy, we set $\xi_{b,1}$, $\xi_{b,2}$, and $\xi_{b,3}$ to 0.9, 0.32, and 0.15, respectively; and b_1, b_2, and b_3 equal 0.21, 1.29, and 2.42 times average earnings in the economy.[26]

Government debt and expenditure. The amount of government debt B is set to target a debt-to-output ratio of 60%, resulting in $B = 38.09$. The stationary government period budget constraint, equation (18), then implies that government expenditures on public goods, as a fraction of GDP is $G/Y = 13.3\%$.

4.2.5 Initial Assets and Borrowing Limits

The distribution of household initial assets when entering the economy is calibrated from the PSID. We approximate the empirical distribution of net worth of young (age-25) households by a discrete distribution with fifty mass points, each representing 2% of the population. The initial assets of newborn households in the model are then drawn randomly from this discrete distribution.[27] Figure 21 in Appendix C shows that the discrete distribution provides a close approximation to its empirical counterpart. To be consistent with the empirical distribution of initial assets, the borrowing limit at the beginning of the life cycle is set to the lowest mass point of the discretized distribution of initial assets, corresponding to −\$92,283 per adult in 2016 dollars. The borrowing limit is tightened gradually as households age such that it reaches zero at retirement.[28]

---

[26] In other words, the marginal replacement rate is 90% up to 0.21 times the average earnings in the economy, 32% between 0.21 and 1.29 times the average earnings, 15% between 1.29 and 2.42 times the average earnings, and 0% above.

[27] The correlation between initial net worth and log-wages of young households is close to zero in the data; therefore, initial assets are assumed to be independent of initial labor productivity in the model.

[28] Figure 22 of Appendix C presents the share of borrowing constrained households in the calibrated model.

Table 1: Model Parameters

Parameter	Governing	Value
A. Demographics		
(J_R, J)	retirement age and end of life cycle	$(42, 76)$
B. Preferences		
(σ, η, β)	risk aversion, labor elasticity, and discount factor	$(1.5, 0.5, 0.991)$
$(\bar{\psi}, \sigma^2_\psi)$	level and dispersion of labor disutility	$(2.116, 0.760)$
C. Taxes		
(μ, τ)	income tax progressivity and level	$(0.137, 0.105)$
$(\tau_c, \tau_k, \tau_{ss})$	consumption, capital, and payroll taxes	$(0.041, 0.330, 0.124)$
D. Retirement Benefit		
(b_1, b_2, b_3)	cutoff levels of average earnings	$(0.21, 1.29, 2.42)$
$(\xi_{b,1}, \xi_{b,2}, \xi_{b,3})$	marginal replacement rates	$(0.90, 0.32, 0.15)$
E. Factor Prices and Technology		
(r, w)	interest rate and price of effective labor	$(3\%, 1)$
(Z, ζ, δ)	TFP, capital share, and depreciation rate	$(0.985, 0.330, 0.110)$
F. Others		
(B, G)	government debt and expenditure	$(38.09, 8.47)$

4.2.6 Production Technology and Factor Prices

We consider two functional forms for the constant-returns-to-scale production function $F(K, N)$. As baseline, we assume a linear production technology:

$$F(K, N) = [(r + \delta)K + wN]. \tag{19}$$

This formulation implies that the interest rate and the wage per efficiency unit of labor (r, w) are invariant to changes in tax policy. Thus, our baseline analysis considers a partial equilibrium framework with fixed factor prices.[29] The interest rate r is set at 3%, and the price of effective labor w is normalized to one. We also conduct a general equilibrium analysis in Section 6.3 where the production function takes a Cobb-Douglas form:

$$F(K, N) = ZK^{\zeta}N^{1-\zeta}. \tag{20}$$

We set the capital share ζ to 0.33. Total factor productivity Z is calibrated so that the equilibrium price of effective labor w is exactly one (a normalization). The capital depreciation rate δ is chosen to ensure an equilibrium interest rate of $r = 3\%$. Table 1 summarizes the calibration of the model and the values of model parameters.

---

[29] Alternatively, we can interpret the benchmark as an open economy facing a fixed global interest rate and free capital mobility. Under this interpretation, the fixed interest rate implies a fixed wage due to the constant-returns-to-scale assumption.

4.3 Goodness of Model Fit

We now examine the model's performance in explaining the data by comparing household life-cycle profiles from the model with those estimated from the PSID data.[30] Our main conclusion is that the model closely replicates the life-cycle profiles of key household variables and their dispersion, despite the fact these profiles were not targeted in the calibration.

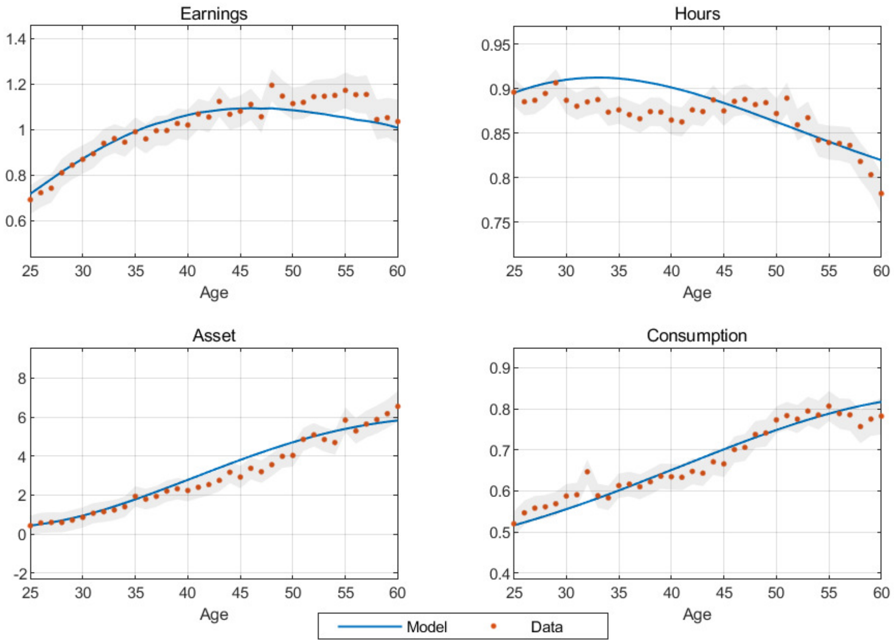

Figure 1: Model vs. Data: Life Cycles of Household Averages

Notes: This figure shows the life cycles of cross-sectional means in the benchmark model (blue solid lines) and in the PSID data (red dotted lines) together with the 95 percent confidence interval (grey bands). The consumption life cycle from the data is scaled up to match the life cycle average of consumption in the model.

Figure 1 displays the life-cycle profiles of average earnings, hours worked, assets, and consumption from age 25 to 60, for the model (solid blue lines) and the PSID data (red dotted lines with shaded 95 percent confidence intervals). In the data, average earnings show rapid growth among young households, stabilize during mid-life, and decrease near retirement. Average hours worked remain relatively stable until age 50, after which they decline notably. Average asset consistently grows until age 60, whereas average consumption shows a similar growth pattern. The model matches these empirical life-cycle profiles well.[31]

---

[30] When estimating the life-cycle profiles from the PSID data, we control for time effects since Heathcote, Storesletten, and Violante (2005) find no evidence for significant cohort effects.

[31] Since consumption data from the PSID do not include all categories of consumption expenditures, in Figure 1, the consumption profile from the data is scaled up by a constant factor such that average consumption in the data is identical to that in the model. The focus of the comparison is hence consumption growth over life cycle. Household consumption data are normalized using the OECD-modified equivalence scale as in Kaplan (2012). Whereas raw consumption data exhibit a hump-shaped life-cycle profile, this pattern largely disappears after adjustment using the equivalence scale, which removes the effects of changes

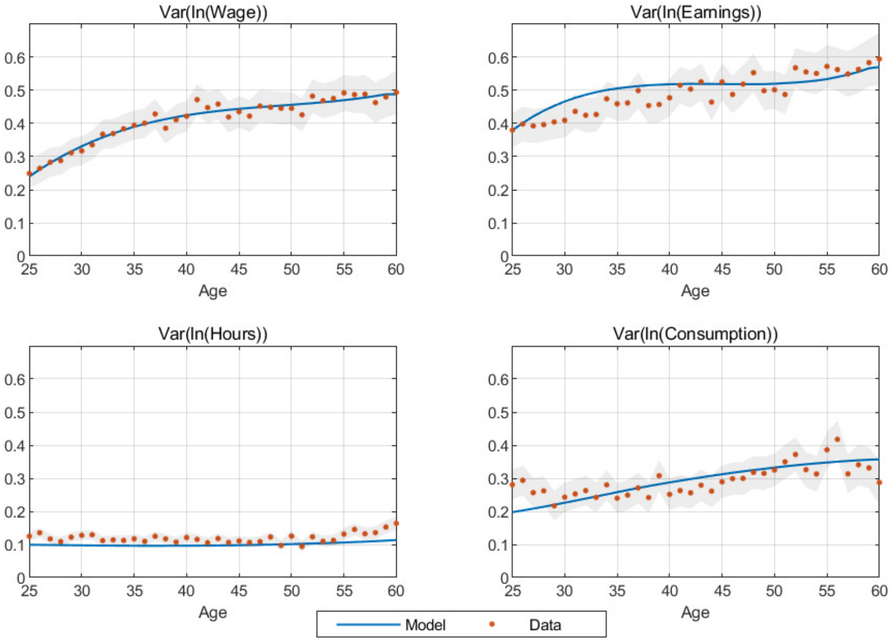

Figure 2: Model vs. Data: Life Cycles of Inequality

Notes: This figure shows the life cycles of cross-sectional variances in the benchmark model (blue solid lines) and in the PSID data (red dotted lines) together with the 95 percent confidence interval (grey bands). The life cycle of log-consumption variance in the model is shifted to match the life cycle average in the data.

Figure 2 presents the life-cycle profiles of wage, earnings, hours worked, and consumption inequality, measured by the cross-sectional variance of logs at each age, plotted on the same scale for ease of comparison. In the data, wage inequality rises from 0.25 at age 25 to 0.49 at age 60. Inequality in hours worked is much smaller in comparison, and remains relatively constant between age 25 and 60. As a result, the variance of log-earnings closely mirrors the life-cycle pattern of wage inequality, increasing from 0.38 to 0.60 between age 25 and 60. Consumption inequality in the data exhibits a gradual upward trend with age, resulting in a roughly 0.1 increase over life cycle, a pattern that is also reflected in the model.[32]

The focus of the paper is on taxation over the life cycle, and therefore so far we have focused on this dimension for the discussion of model validation. As is well recognized in the literature, standard incomplete-markets life-cycle models without entrepreneurship provide a reasonable representation of the bottom of the wealth distribution, but struggle to match the very high degree of wealth concentration at the very top of the distribution observed in

---

in household composition, especially variation in the number of adults and children over the life cycle.

[32] Given the fact that consumption data is subject to large measurement error, we do not attempt to match the overall level of consumption dispersion in the data. Hence, in Figure 2, the life-cycle variance profile of log consumption from the model is shifted such that, *on average*, the model variances match their counterparts in the data. The key question for model validation therefore is whether the model implies empirically plausible *changes* in the variance over the life cycle.

the data, see e.g., Quadrini (1999, 2000) or Cagetti and De Nardi (2006). Our model shares this property, as shown in Table 11 of Appendix C. Although the model aligns well with the data up to the 60th percentile of the distribution, the Gini coefficient for net worth is noticeably lower in the model at 0.61, relative to its value of 0.82 in the data (PSID, from 1999–2017). This shortcoming is mostly due to the fact that our model predicts too low a share of wealth by households at the very top of the distribution.

The key role of wealth in our model is to provide consumption insurance, in addition to funding retirement consumption. Given our focus on redistribution and insurance through progressive income taxation, what mostly matters for our model to be a credible laboratory for this question is that the model reproduces the degree of consumption insurance observed in the data. In Appendix C we use the methodology introduced by Blundell, Pistaferri, and Preston (2008) to measure the degree of consumption insurance against income shocks in the model. Table 12 shows that the degree of consumption insurance implied by our model aligns well with the existing empirical evidence. The reason for this consistency is that we calibrate the model to match the average level of net worth in the data, and the model also replicates the life-cycle profile of average net worth well (and the degree of consumption insurance is fairly invariant to net worth at the very top of the wealth distribution).

5 Economic Consequences of a Lifetime Income Tax

We now compare the positive and welfare properties of a lifetime income tax and an annual income tax. We first quantify the welfare gains of social welfare-maximizing AIT and LIT reforms. We then investigate the sources of the welfare gains of a LIT, emphasizing improved labor efficiency through reallocation of household hours. Lastly, we discuss the life-cycle implications and transitional dynamics induced by potential tax reforms.

5.1 The Optimal Tax Problem

The government (both under a progressive LIT and AIT) faces a classic equity-efficiency trade-off. On one hand, a progressive income tax provides public insurance against idiosyncratic wage risk and redistribution against ex ante heterogeneity. On the other hand, it induces efficiency losses by distorting household labor supply. Optimal policy must balance these gains and losses. As explained in Section 3.3.1, income tax policy is summarized by the tax function $\widetilde{T}(y, Y)$ that specifies household tax liability in the current period. Depending on the type of income tax, it takes the form:

$$\widetilde{T}(y, Y) = \begin{cases} T(y), & \text{if annual income tax;} \\ T(Y+y) - T(Y), & \text{if lifetime income tax.} \end{cases}$$

Here y denotes current-year earnings, Y is accumulated lifetime earnings prior to the current year, and $T(\cdot)$ controls the mapping from pre-tax income to after-tax income. As in the simple model the tax function $T(\cdot)$ is given by equation (2), with parameters (μ, τ) controlling the progressivity and the level of tax rates, respectively.

We assume that policymakers seek to maximize the expected lifetime utility of a newborn cohort in the stationary equilibrium by choosing the income tax policy, as represented by the policy parameters μ and τ. The optimal income tax problem can then be written as

$$\max_{(\mu,\tau)} \int V(\mathbf{S}; \mu, \tau) d\Phi_1(\mathbf{S}),$$

subject to the government budget constraint; $V(\mathbf{S}; \mu, \tau)$ is the value function of state-$\mathbf{S}$ household under tax policy (μ, τ), and $\Phi_1(\mathbf{S})$ is the measure of age-1 newborn, state-$\mathbf{S}$ households.[33]

Since our goal is to study the potential of a LIT for providing insurance and redistribution against idiosyncratic risk, we require government policy to satisfy a within-cohort budget constraint to avoid intergenerational transfers through the income tax system. The within-cohort budget constraint states that the age-1 value of total government revenues collected from each cohort, including initial net worth (as negative revenues) and accidental bequests, discounted by the market interest rate r, must be the same as under the status quo policy.[34]

Note that in a LIT, the discount rate for computing lifetime income is, in principle, also a policy parameter that the government may choose. As baseline, we set this rate to zero which also makes the LIT function $\widetilde{T}(y, Y)$ age-independent, and thus simpler.[35] In Section 6.1, we show that our main findings are robust to alternative choices of this discount rate.

Our baseline analysis also assumes that the interest rate and hence the wage rate of effective labor are fixed at their status quo levels and thus policy invariant, as in a small open economy. In Section 6.3, we consider a closed economy in general equilibrium in which the interest rate and wage respond fully to changes in the tax system.

---

[33] We also require that, when changing tax policy, the ratio of initial net worth of newborns to accidental bequests remains the same as in the benchmark, achieved by scaling initial net worth distribution accordingly.

[34] Appendix C.7 shows that if tax reforms are restricted to satisfy the within-period government budget constraint with constant government expenditure and debt (rather than insisting on constant tax revenue being collected from a newborn cohort), then newborn welfare-maximizing (in stationary equilibrium) tax reforms lead to much larger welfare gains from a LIT. However, these gains largely stem from redistributing welfare from old to newborn generations. An optimal policy analysis would then require taking a stance on how to weigh different generations, which we avoid by imposing the within-cohort budget constraint.

[35] Note that the choice of this discount rate does not affect how the government discounts tax revenues in its within-cohort budget constraint, which always uses the market interest rate r. Also, since households earnings are subject to large idiosyncratic risks, the risk-free interest rate r is not necessarily the proper choice for measuring lifetime income. See Huggett and Kaplan (2011) for more discussion on this issue.

5.2 Welfare Gains of a Lifetime Income Tax

Does a LIT achieve higher social welfare than an AIT? To address this question, we first examine the welfare effect of replacing the status quo AIT with a LIT that maintains the same progressivity and tax revenue. We then allow tax progressivity to adjust and compare maximized welfare under the optimal annual and lifetime income tax, as determined by the tax problem in Section 5.1. Table 2 presents the four tax policies considered and their associated welfare gains relative to the status quo.

Table 2: Welfare Gains from Annual and Lifetime Income Taxes

	Annual Tax		Lifetime Tax	
	Status Quo	Optimal	Status Quo μ	Optimal
Progressivity (μ)	0.137	0.194	0.137	0.242
Level (τ)	0.105	0.110	-0.431	-1.035
Level (comparable)	10.5%	11.0%	14.3%	17.7%
Avg. Tax Rate	13.8%	14.9%	15.7%	19.3%
Welfare Gain	–	0.19%	0.63%	1.27%

Notes: Welfare changes are in consumption equivalent variations as percentages of household lifetime consumption in the status quo. “Level (comparable)” is the average tax rate of a household with constant earnings equal to the status quo average in each working year. “Avg. Tax Rate” is the ratio between total labor income taxes and total labor earnings in the economy.

The third column of Table 2 shows that even if the tax progressivity parameter μ remains at its status quo level, switching to a LIT that collects the same tax revenue already leads to a sizable welfare gain of 0.63% of household lifetime consumption. Note that the tax progressivity parameter μ is comparable between an AIT and a LIT since $1 - \mu$ represents the pass-through rate from pre-tax to after-tax (annual or lifetime) earnings. On the other hand, the tax level parameter τ is not directly comparable between the two tax systems. Therefore, to compare the tax levels, we use the average tax rate faced by a hypothetical household with constant earnings (equal to average earnings in the status quo economy) during all working years, labeled as “Level (comparable)”, which is 14.3% under the LIT, compared to 10.5% under the status quo policy. We also report the average tax rate, defined as the ratio between total labor income tax receipts and aggregate labor income.

When tax progressivity is optimally chosen, the advantage of a LIT becomes even more pronounced, yielding additional welfare gains of 1.08% of household lifetime consumption, (comparing the second and fourth columns of Table 2). Under the AIT, the optimal policy is substantially more progressive than the status quo, as indicated by the increase in tax progressivity μ from 0.137 to 0.194. This means that a 10% increase in pre-tax earnings leads to only a 8.06% increase in after-tax earnings under the optimal policy, compared to

8.63% under the status quo. However, the welfare gain from higher progressivty under the AIT remains modest at 0.19% of lifetime consumption. In comparison, the optimal LIT is more progressive than the optimal AIT (0.242 vs. 0.194) and also achieves substantially larger welfare gains, 1.27% of lifetime consumption.

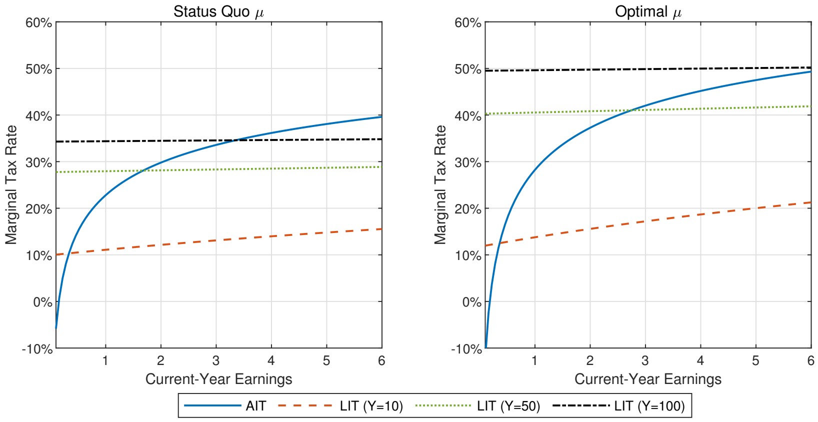

Figure 3: Marginal Tax Rate Under Annual and Lifetime Income Taxes

Notes: This figure plots the marginal tax rate on current-year earnings under AIT (blue solid lines) and LIT with accumulated lifetime earnings equal to 10, 50, and 100 income units (red dashed lines, green dotted lines, and black dash-dotted lines). In the left panel, tax progressivity μ is fixed at its status quo level 0.137, whereas the right panel refers to the optimal annual and lifetime income taxes.

Figure 3 shows how marginal tax rates vary with current-year earnings under the AIT (blue solid line) and the LIT, at different levels of accumulated lifetime earnings (red dashed, green dotted, and black dash-dotted lines). The left panel corresponds to status quo tax progressivity, and the right panel to optimal progressivity. Marginal rates rise fast with current-year earnings under the AIT, more so under the *optimal* (more progressive than the benchmark) AIT. In contrast, under the LIT, marginal tax rates are almost constant in current-year earnings except when accumulated lifetime earnings are low, and the progressivity of the LIT is more strongly reflected by the rising marginal rates with accumulated earnings. Since the optimal LIT is more progressive than the status quo, the gaps in marginal tax rates between different levels of accumulated earnings are larger in the right panel.

5.3 Understanding the Welfare Gains

5.3.1 Aggregate and Distributional Implications

To provide intuition for the results in Table 2, Table 3 presents changes in aggregate variables and inequality (variance of logs) resulting from tax reforms replacing the status quo AIT with the optimal AIT (first column), the LIT with status quo progressivity (second column), and the optimal LIT (third column). Our analysis in Section 2 suggests that the primary advantage of a LIT is that it promotes a more efficient distribution of labor supply over time and across states, but at the cost of worse consumption insurance later in the life cycle compared to an AIT system. The results of the quantitative model confirm this insight, as shown in the second column of Table 3. When switching from the status quo AIT to a LIT while maintaining the same tax progressivity, total hours decline by 0.99%, whereas total earnings rise by 0.23%. Aggregate labor efficiency, again defined as total earnings divided by total hours worked, rises by 1.23%. This occurs because, relative to the status quo, households work longer hours when their wages are high, as reflected in increased hours and earnings inequality. On the other hand, switching to the LIT also reduces consumption insurance later in life and amplifies the increase in consumption inequality over the life cycle, with inequality lower at younger ages and higher at older ages.[36] Overall consumption inequality remains nearly unchanged. Ultimately, the benefits of improved labor efficiency outweigh the costs of worsened consumption insurance, and switching to a LIT leads to a sizable welfare gain equivalent to 0.63% of lifetime consumption.

Table 3: Aggregate and Distributional Implications of Tax Reform

	Annual Tax	Lifetime Tax	
	Optimal	Status Quo μ	Optimal
Hours	−1.78%	−0.99%	−5.06%
Earnings	−1.97%	0.23%	−3.44%
Consumption	−3.32%	0.38%	−5.24%
Labor Efficiency	−0.19%	1.23%	1.71%
Hours Inequality	−2.99%	12.48%	14.80%
Earnings Inequality	−1.35%	8.05%	10.81%
Consumption Inequality	−11.80%	−0.07%	−20.45%

Notes: This table reports the changes in aggregate variables and inequality as percentages of their status quo levels, induced by tax reforms to the policy of each column. Inequality is measured by variance of logs.

When tax progressivity is chosen to maximize welfare, as in Section 2 for the simple model, the main trade-off is between a more efficient distribution of consumption and a higher level of labor supply relative to consumption. A more progressive income tax improves consumption insurance and redistribution but depresses overall labor supply relative to consumption. Under the AIT system (first column), the optimal policy is more progressive than the status quo, and thus reduces consumption inequality substantially by 11.80% but also lowers total hours and earnings by 1.78% and 1.97%, respectively. Since total earnings fall more than total hours, labor efficiency declines by 0.19%.

In contrast, under the LIT system (third column), greater tax progressivity improves labor efficiency, providing an additional benefit of more progressive income taxation. Consequently, the optimal LIT is more progressive than the optimal AIT. It lowers hours worked

---

[36] This pattern is illustrated in the bottom right panel of Figure 6 below.

and earnings further, but also reduces consumption inequality more significantly by 20.45% and improves labor efficiency by 1.71%. As a result, it achieves a higher level of welfare than the optimal AIT, with the difference amounting to about 1.08% of lifetime consumption.

Why does labor efficiency rise with tax progressivity under the LIT but falls under the AIT in the calibrated version of our model? First, when the income tax becomes more progressive, consumption inequality falls, narrowing the differences in marginal utility of consumption between high- and low-wage states. This shift in the distribution of marginal utility of consumption induces households to work more in high-wage states as their relative consumption in these states is lower under a more progressive tax policy, leading to an improvement in average labor efficiency. However, there is a second effect of a more progressive income tax: marginal tax rates rise in high-wage states and fall in low-wage states. This shift incentivizes households to allocate labor supply less strongly toward high-wage times and states than is efficient. This second effect is stronger under the AIT than under the LIT because marginal tax rates are more sensitive to current earnings under the AIT. In the absence of income effects on labor supply ($\sigma = 0$) we demonstrated this point theoretically in Corollaries 4 of the simple model in Section 2.4, and it remains true with $\sigma > 0$ in the calibrated quantitative model. As a result, a more progressive LIT tends to enhance labor efficiency whereas it reduces labor efficiency with a more progressive AIT, at least as long as the substitution effect is not too potent relative to the income effect (in which case aggregate labor efficiency might decline with higher tax progressivity even under the LIT).[37]

5.3.2 Decomposition of Welfare Gains

Our previous discussion suggests that one key difference between the annual and lifetime income taxes is their divergent effects on labor efficiency. The LIT improves labor efficiency relative to the AIT, and this efficiency gain increases with the progressivity of the LIT. As a result, policymakers prefer a more progressive tax system with the LIT and achieve larger welfare gains (Table 2). The more progressive LIT reduces total hours and consumption as well as consumption inequality, but raises hours inequality (Table 3).

To quantify the contributions of these different channels, we now conduct a welfare decomposition following the method in Conesa et al. (2009), which separates welfare gains from changes in the average levels of consumption and labor and their distributions. As in Moschini and Tran-Xuan (2025), we further decompose the distributional effects into changes in the age profiles of consumption and labor and the remaining distributional changes.[38] We

---

[37] What is true in our calibrated model with $\sigma = 1.5$, and we believe is true in general (based on the theory in Section 2 and model simulations) is that for the separable utility function used (for which income and substitution effects are controlled separately by σ, η), aggregate labor efficiency moves either more positively or less adversely with tax progressivity under a LIT than an AIT.

[38] The age-profile effects capture the welfare impact of reshuffling consumption or labor over the life cycle

consider all possible orders of introducing these changes and compute the weighted averages of the decomposition results, following the Shapley-Owen-Shorrocks scheme outlined in Audoly, McGee, Ocampo-Diaz, and Paz-Pardo (2025), so that the results are order-independent and additively separable. Table 4 contains the results.[39]

Table 4: Decomposition of Welfare Gains

	Annual Tax	Lifetime Tax	
	Optimal	Status Quo μ	Optimal
Total Welfare Gain	0.19%	0.63%	1.27%
Consumption	−1.46%	0.53%	−1.95%
Level	−3.39%	0.38%	−5.46%
Age	0.35%	0.15%	0.75%
Distribution	1.59%	−0.01%	2.76%
Labor	1.65%	0.10%	3.22%
Level	1.38%	0.79%	3.91%
Age	0.06%	−0.12%	−0.12%
Distribution	0.20%	−0.58%	−0.56%

Notes: Details of the decomposition method are in Appendix C.2.

For the optimal AIT (first column), the total welfare gain of 0.19% is the combined result of welfare losses from worse consumption (−1.46%) and welfare gains from labor (1.65%). The welfare effect from consumption can be further decomposed into the impact from a change in average consumption ("Level"), the change in the age-profile of consumption ("Age"), and the remaining change in the distribution of consumption ("Distribution"). The optimal policy attains a more equal distribution of consumption, improving welfare by 1.59%. However, the loss from lower average consumption (−3.39%) dominates this gain from reduced consumption dispersion; the net welfare effect from changes in consumption is negative. The change in the age-profile of consumption generates a small welfare gain (0.35%), suggesting that households can better align consumption with their time preferences.

For the welfare effect through household labor, we similarly separate it into effects from the change in average hours worked, the change in the age-profile of hours, and the remaining distributional changes. The optimal policy reduces average hours worked, and households enjoy more leisure which improves welfare by 1.38%. The change in the distribution of hours also improves welfare by 0.20% because household disutility from labor is convex and hours inequality falls under the optimal AIT. The labor age effect is slightly positive (0.06%), indicating that hours are more closely aligned with time preferences.

---

while holding total consumption or labor unchanged. Roughly speaking, a positive effect indicates that the discounted marginal utility (or disutility) of consumption (or labor) becomes more equal over the life cycle.

[39] Appendix C.2 provides details on the decomposition method. Table 13 also presents decomposition results with the order of changes in Conesa et al. (2009) which are very similar to the weighted results.

In comparison, most of the 0.63% welfare gain from a switch to the LIT with status quo progressivity (second column) stems from rising average consumption (0.38%). The welfare gain from labor is positive but small (0.10%) because the welfare gain from more leisure (0.79%) and the welfare loss from a more dispersed hours distribution (−0.58%) offset each other. The consumption age effect is positive (0.15%), whereas the labor age effect is negative (−0.12%), indicating that hours become more aligned with market prices (wages and interest rate) rather than time preferences. Raising tax progressivity of the LIT to its optimal level (third column), households enjoy more leisure and suffer less consumption inequality at the cost of lower average consumption. The net welfare effect from consumption is close to that associated with the optimal AIT, and the extra 1% total welfare gain from the optimal LIT is mostly due to the incremental welfare gain from increased leisure.

5.3.3 Insurance vs. Redistribution

A progressive income tax system provides not only insurance against ex post earnings risk (persistent and transitory wage shocks) but also redistribution against ex ante heterogeneity (initial labor productivity, the level of labor disutility, and initial wealth). To separate the role of the LIT in providing insurance and redistribution, we conduct counterfactuals in which we shut down ex ante heterogeneity or ex post earnings risks.[40] For comparison, the first two columns of Table 5 reproduce the optimal annual and lifetime income tax policies and their associated welfare gains in the economy with both types of household heterogeneity.

Table 5: Insurance vs. Redistribution

	All Heterogeneity		Only Earnings Risks		Only Ex Ante Heterogeneity	
	Annual Tax	Lifetime Tax	Annual Tax	Lifetime Tax	Annual Tax	Lifetime Tax
Progressivity (μ)	0.194	0.242	0.101	0.161	0.238	0.238
Level (τ)	0.110	−1.035	0.101	−0.552	0.116	−1.039
Level (comparable)	11.0%	17.7%	10.1%	14.8%	11.6%	16.4%
Avg. Tax Rate	14.9%	19.3%	11.8%	15.2%	11.6%	15.5%
Welfare Gain	0.19%	1.27%	0.06%	0.65%	0.52%	0.65%

The third and fourth columns report the corresponding results in the counterfactual economy with only earnings risks (i.e., without ex ante heterogeneity). The optimal annual and lifetime income taxes are still progressive, as indicated by the positive tax progressivity μ, but the degrees of progressivity are lower than in the full model. This is not surprising because: i) there is no redistribution motive anymore; and ii) wage shocks are either moderately persistent or completely transitory and can be insured reasonably well through precautionary savings. The welfare gains from the optimal annual and lifetime income taxes relative to the

---

[40] Ex ante heterogeneity is eliminated by assigning newborns average initial assets, setting the dispersion of labor disutility to zero, and fixing initial labor productivity at the median. Ex post earnings risk is removed by holding each household's idiosyncratic labor productivity fixed at its initial draw over the life cycle.

counterfactual status quo economy are 0.06% and 0.65% of lifetime consumption, respectively. The optimal LIT is still more progressive than the optimal AIT and achieves larger welfare gains since the previous labor efficiency argument still applies.

The last two columns display the results with only ex ante heterogeneity. Since ex ante heterogeneity is permanent, its welfare effects are more difficult to moderate by household choices, giving rise to a motive for government intervention. Consequently, the optimal tax polices are considerably more progressive relative to the case with only earnings risks. The welfare gains are 0.52% and 0.65% of lifetime consumption for the optimal AIT and LIT, respectively, which stem mostly from reduced consumption inequality and increased leisure under a more progressive tax system. Since without the earnings risks there is less variation in labor productivity over time and across states, the labor efficiency channel becomes weaker, and the optimal progressivity is roughly the same across the AIT and the LIT. Overall, Table 5 suggests that the advantage of the LIT over the AIT is most prominent when both ex-post earnings risks and ex-ante heterogeneity are present.

5.4 Life-Cycle Profiles

Since the LIT differs strongly from the AIT in the way it collects taxes over the life cycle, it is instructive to examine, in this section, how the LIT impacts life cycle choices.

5.4.1 Life Cycles of Household Averages

Figure 4 presents life-cycle profiles of average hours worked, earnings, after-tax earnings, and asset holdings under four regimes: the status quo AIT (blue solid lines), the optimal AIT (red dashed lines), the LIT with the status quo progressivity (black dash-dotted lines), and the optimal LIT (green dotted lines). Figure 5 does the same for tax rates.

The life-cycle profiles of hours worked (top-left panel) are shaped by the degree of patience, the life-cycle wage profile, and tax policy. A high degree of patience (i.e., $\beta(1+(1-\tau_k)r)>1$) implies that labor supply should fall with age; thus the overall downward trends in the graph. However, since the deterministic life-cycle wage profile is hump-shaped (Figure 11 in Appendix C), young households also want to delay labor supply until wages are higher. As a result of these two forces, hours worked tend to peak earlier than wage before declining.

Under the LIT, the average marginal tax rate increases sharply with age (top-right panel of Figure 5). Consequently, switching to a LIT while maintaining status quo progressivity reduces hours worked among older households, since higher tax rates at older ages further discourage labor supply. For younger households, labor supply increases due to negative marginal tax rates early in life under the LIT. However, these effects are quantitatively less significant, as their labor supply decisions are also influenced by the understanding that higher current earnings will lead to higher future tax liabilities under the progressive LIT.

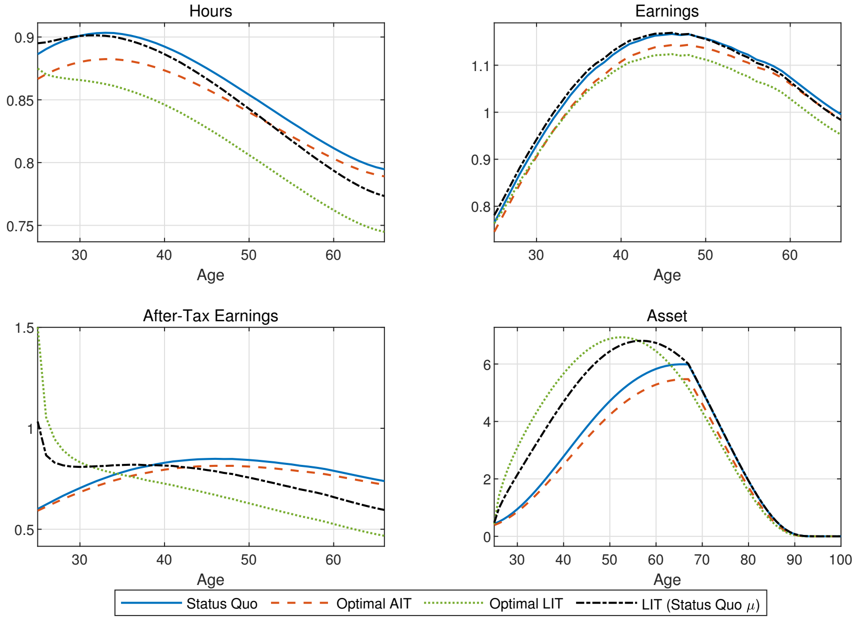

Figure 4: Life-Cycle Profiles of Household Averages

Notes: This figure plots the life-cycle means in the status quo economy (blue solid lines), under the optimal annual income tax (red dashed lines), under the optimal lifetime income tax (green dotted lines), and under the lifetime income tax with the status quo tax progressivity (black dash-dotted lines).

Under both the AIT and the LIT, increasing tax progressivity to the optimal level depresses labor supply at all ages, shifting the entire life-cycle hours profile downward.

Household earnings in the top-right panel of Figure 4 largely follow the hump-shaped life-cycle wage profile, but with an earlier peak due to declining hours with age. Switching to the LIT (with status quo progressivity) leads to slightly higher earnings of young households and lower earnings of old households; moving to the optimal annual or lifetime income tax lowers earnings at all ages. The bottom-left panel shows the life-cycle profiles of after-tax earnings, which differ significantly between the AIT and the LIT. With a AIT, after-tax earnings closely track household earnings. In contrast, under a progressive LIT, households receive transfers (negative taxes) from the government early in life when accumulated lifetime earnings are low (top-left panel of Figure 5). As accumulated lifetime earnings rise with age, taxes increase, causing after-tax earnings to decline.[41] As the bottom-right panel shows, under the AIT, households accumulate wealth gradually over their working years to fund retirement consumption, but also for precautionary reasons to hedge against stochastic wage fluctuations. Household asset holdings peak near retirement as is common in life-cycle models. In contrast, under the LIT, young households accumulate substantially more assets, and household wealth peaks earlier. Households receive transfers when young and pay higher

---

[41] Redistributing income toward young households can improve welfare by relaxing borrowing constraints. However, as shown in Figure 22 of Appendix C, only a small share of households are borrowing-constrained before retirement. Therefore, this is not a main reason for the larger welfare gains from the LIT.

taxes later as accumulated lifetime earnings rise with age. Thus, they save most of these transfers to finance future taxes, explaining the rapid accumulation of wealth early in life.

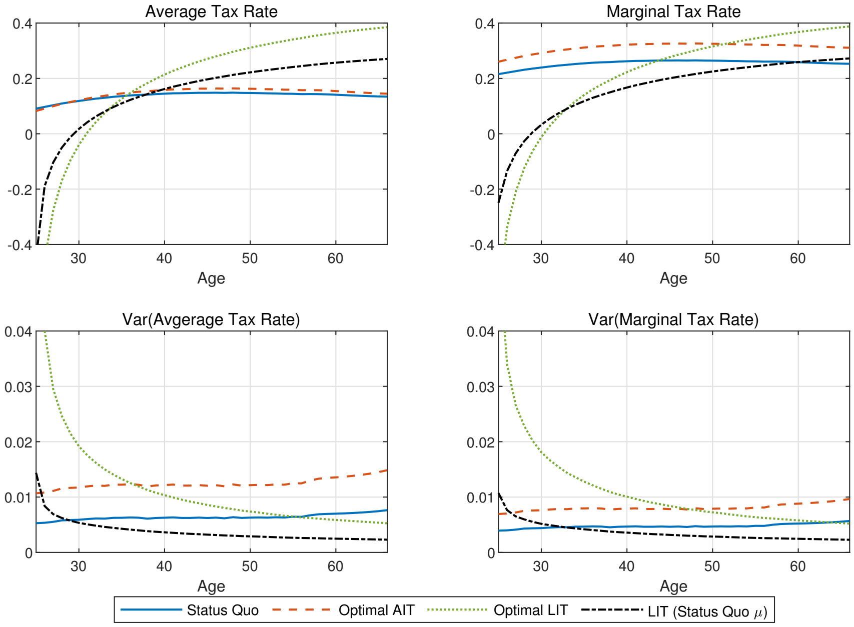

Figure 5: Life-Cycle Profiles of Average and Marginal Tax Rates

Notes: This figure plots the life-cycle profiles of cross-sectional income-weighted means (top panels) and variances (bottom panels) of average (left panels) and marginal (right panels) tax rates in the status quo (blue solid lines), under the optimal AIT (red dashed lines), under the optimal LIT (green dotted lines), and under the LIT with status quo tax progressivity (black dash-dotted lines).

5.4.2 Life Cycles of Labor Efficiency and Inequality

The top-left panel of Figure 6 shows life-cycle profiles of labor efficiency at each age divided by its counterpart in the status quo, under the four tax regimes. The optimal AIT reduces labor efficiency at all ages compared to the status quo, especially for households in their 30s. In contrast, the LIT improves labor efficiency throughout the life cycle, and the efficiency gains rise with age, more so under the optimal LIT. This is consistent with our main message from Section 2 that a LIT improves labor efficiency through a more efficient distribution of hours worked across contingent wage states. As wage dispersion grows with age due to the accumulation of shocks, so do labor efficiency gains. In addition, since this efficiency gain is enhanced by the progressivity of the LIT, it rises more over the life cycle under the optimally more progressive LIT.

The remaining panels of Figure 6 display how hours, earnings, and consumption inequality (measured as variance of logs) evolve over the life cycle. Hours inequality (top-right panel) and earnings inequality (bottom-left panel) rise with age due to the accumulation of idiosyncratic wage shocks. Compared to the status quo, the optimal AIT reduces hours and earnings inequality at all ages, whereas the LIT raises them substantially throughout life.

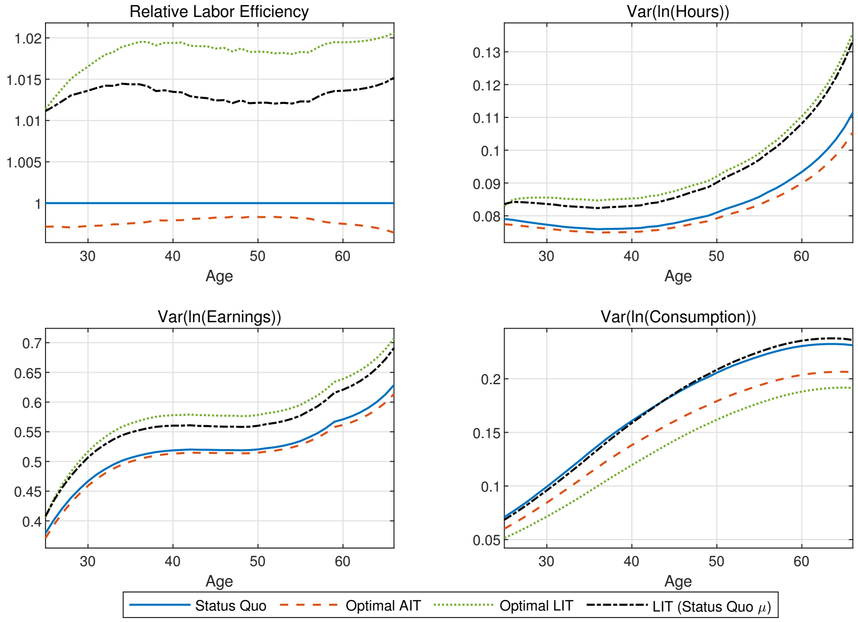

Figure 6: Life-Cycle Profiles of Labor Efficiency and Inequality

Notes: This figure plots the life-cycle profiles of relative labor efficiency and cross-sectional variances in the status quo (blue solid lines), under the optimal AIT (red dashed lines), under the optimal LIT (green dotted lines), and under the LIT with the status quo tax progressivity (black dash-dotted lines).

Most of these differences are due to the shift from annual to lifetime income taxation; raising progressivity of the LIT further to its optimal level has small effects. Consumption inequality (bottom-right panel) also increases over the life cycle as the accumulation of idiosyncratic wage shocks leads to greater dispersion within each cohort. As in Section 2, switching to the LIT with the same progressivity as the status quo results in higher consumption inequality later in life among old households, although the increase is small. Both the optimal AIT and LIT are more progressive than the status quo, thereby reducing consumption inequality significantly. The reduction is more pronounced under the optimal LIT as it is more progressive than the optimal AIT.

The bottom panels of Figure 5 show that the dispersion of average and marginal tax rates increase with age under the AIT, reflecting growing within-cohort inequality in wages and earnings due to the accumulation of idiosyncratic shocks. Under the LIT, however, the dispersion is higher early in life and declines with age, since with the HSV tax function we adopted, tax rates are more sensitive to cumulative earnings when those earnings, and consequently tax rates, are low. A more progressive tax system, whether annual or lifetime, generally results in greater dispersion of tax rates across individuals at all ages.

5.5 Transition to the Optimal Stationary Equilibrium

The optimal policies reported in Table 2 maximize social welfare in stationary equilibrium, and the LIT outperforms the AIT in terms of maximum welfare gain by about 1%

of household lifetime consumption. However, what are the welfare consequences of these reforms for households during the transition from the status quo economy to the optimal stationary equilibrium? To understand these, we compute the transitions induced by the optimal tax reforms in Table 2. Starting from a stationary equilibrium with status quo AIT policy, at time $t = 0$, a permanent change in income tax policy occurs (i.e., switching to the optimal annual or lifetime income tax), and the economy evolves endogenously towards that stationary equilibrium. We label different generations of households by the time of their birth; generation x is born at time $t = x$. At the time of the tax reform, we consider two scenarios: i) the new policy is only applicable to new households (i.e., current and future new-born households), and households born before the reform are still subject to the status quo policy; or ii) all current and future households are subject to the new tax policy.[42]

As mentioned in Section 5.1, to focus on the insurance and redistribution role of income taxes, we impose a within-cohort government budget constraint when solving for the optimal policy. This guarantees that the present-value of total tax revenues collected from each cohort is the same between the status quo and the optimal policies. However, this also implies that the amount of government debt must differ between the status quo and the optimal stationary equilibrium to simultaneously satisfy the government period budget constraint. In Appendix C.3, we show that if the tax reform is only applicable to new households, and the within-cohort government budget constraint is satisfied, then the level of government debt will evolve endogenously following the government period budget constraint along the transition and eventually converge to a sustainable level in the new stationary equilibrium.

The top-left panel of Figure 7 presents the welfare effects of the AIT and LIT tax reforms for different generations when the new policy is only applicable to new households. Without changes in factor prices (interest rates and wages), current households (generations $x < 0$) are completely insulated from the influence of tax reform. In contrast, new generations immediately enjoy the full benefit of the tax reform, obtaining the welfare gain of the optimal stationary equilibrium, which is larger under the LIT than under the AIT by about 1% of lifetime consumption. Thus, our previous conclusion that the LIT outperforms the AIT remains valid when taking into account transitional dynamics. As the top-right panel shows, with the LIT reform government debt gradually increases during the transition and stabilizes at a higher level, whereas the AIT reform results in a small reduction in public debt.

Suppose now that we subject all current and future households to the new tax policy after the reform. Then the adjustment of government debt alone is no longer sufficient for a

---

[42] In the first scenario, initial net worth of new households is set to its final level, ensuring an immediate transition. In the second scenario, since an immediate transition is not possible, initial new household net worth is set according to the fixed ratio of initial net worth to accidental bequests: in each period, the total initial net worth of newborn households is a constant share of the total bequests from the previous period.

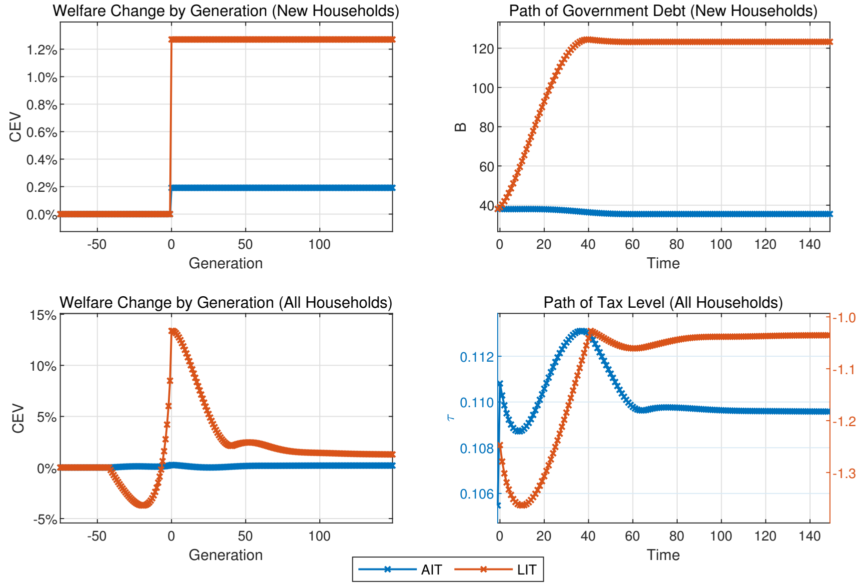

Figure 7: Transition to the Optimal Stationary Equilibrium

Notes: This figure shows the transitions from the status quo to the optimal steady state with annual and lifetime income taxes, assuming that the new policy only applies to new households (top panels) or all households (bottom panels). The left panels display the welfare effects of tax reforms on different generations, the top-right panel plots the paths of government debt, and the bottom-right panel plots the paths of the tax level τ that balance the government budget period-by-period. The tax reform occurs at $t = 0$.

feasible transition towards the optimal stationary equilibrium because the level of government debt would diverge. Therefore, in this scenario, we also allow the tax level τ to vary over time to balance the government period budget constraint along the transition, subject to the restriction that it eventually converges to the tax level of the optimal policy, insuring that the economy still converges to the optimal stationary equilibrium. For comparability, we set the paths of government debt to be the same as those displayed in the top-right panel.[43]

For the AIT reform, the bottom-left panel of Figure 7 shows that current generations (generation $x < 0$) now benefit from the reform but with lower welfare gains compared to future generations.[44] The bottom-right panel shows that the AIT level jumps up immediately after the tax reform, then decreases before rising again and eventually stabilizing at the long-run optimal level. For the LIT reform, the bottom-left panel shows that the current middle-aged and old working generations suffer significant welfare losses from the tax reform, whereas current young generations and generations born within 30 years after the tax reform enjoy even more substantial welfare gains. The level of the LIT gains is low at the start of the transition and gradually rises over time (bottom-right panel).

---

[43] There are in principle infinitely many combinations of the paths of government debt and income tax levels that balance the government period budget constraints along the transition.

[44] The plotted welfare change is the average across all households of the same generation. Therefore, a positive welfare gain does not imply that *all* households within that generation are better off.

6 Robustness Analysis

In this section we probe the robustness of our results with respect to the implementation of the LIT (Section 6.1), the presence of lump-sum transfers (Section 6.2), and general equilibrium effects (Section 6.3). Additional details are contained in Appendices C.4-C.6.

6.1 Implementation of the Lifetime Income Tax

6.1.1 Effects of the Discount Rate for Calculating Lifetime Earnings

To implement a LIT, the government needs to choose a discount rate r_d for calculating lifetime earnings. We now explore the effects of this choice for the performance of the LIT. By assumption, the choice of discount rate r_d affects only the calculation of tax liability in the lifetime income tax function, which in general takes the form

$$\widetilde{T}(y, Y, j) = \left[T\left(Y + \frac{y}{(1+r_d)^{j-1}} \right) - T(Y) \right] (1+r_d)^{j-1},$$

where Y is in present discounted values with discount rate r_d, and the $(1+r_d)^{j-1}$ factor converts the tax that is due into age-j values as it is paid at that age.[45] This discount rate does not affect how households and the government discount future income and revenues in their budget constraints. There are three intuitive choices for r_d: zero, the pre-tax interest rate r, or the after-tax rate $(1-\tau_k)r$. The baseline had $r_d = 0$; in that case, the LIT function $\widetilde{T}(y, Y, j)$ no longer explicitly depends on age. The pre-tax interest rate r is the discount rate of the government, and the after-tax rate $(1-\tau_k)r$ is the private discount rate of households.

Table 6 reports the effects of r_d on the optimal policy choice and the associated welfare gains. For the AIT (first two columns), since r_d does not enter the income tax function, it is nearly irrelevant.[46] For the LIT (last two columns), a higher r_d lowers the welfare gain from the optimal LIT: when r_d increases from zero to the pre-tax interest rate, the welfare gain falls from 1.27% to 0.90% of consumption. Nevertheless, the welfare gains remain larger than in the optimal AIT. Switching to a LIT with status quo progressivity still generates sizable welfare gains compared to the status quo AIT, which remains stable as r_d changes.

A lower discount rate r_d reduces the relative weights on early-life earnings in the calculation of lifetime earnings for tax purposes, leading to lower tax rates for young households.

---

[45] Note that the present discounted value of all tax payments over a household's working life is

$$\sum_{j=1}^{J_R} \frac{\widetilde{T}(y_j, Y, j)}{(1+r_d)^{j-1}} = T\left(\sum_{j=1}^{J_R} \frac{y_j}{(1+r_d)^{j-1}} \right),$$

and it depends only on the present discounted value of the household's realized labor earnings.

[46] The small effect of r_d arises from its role in computing average earnings $\bar{y}$ for retirement benefits, which are implicitly defined by $Y_{J_R} = \sum_{j=1}^{J_R} \frac{y_j}{(1+r_d)^{j-1}} = \sum_{j=1}^{J_R} \frac{\bar{y}}{(1+r_d)^{j-1}}$.

This introduces intertemporal distortions that reduce labor supply of old households. As discussed in Section 2, such distortions may improve consumption insurance by mitigating the impact of wage risks, which are more pronounced later in life due to the accumulation of idiosyncratic wage shocks. In contrast, compared to a zero discount rate, adopting the pre- or after-tax interest rate enhances labor efficiency more. Under the optimal LIT, the consumption insurance channel dominates, explaining why a zero discount rate leads to higher welfare gains than using either the pre- or after-tax interest rates. For the LIT with the status quo progressivity, however, the two effects largely offset each other and the welfare gains are similar across different discount rates.

Table 6: Effects of Discount Rate for Lifetime Earnings (r_d)

		Annual Tax		Lifetime Tax	
Case		Status Quo	Optimal	Status Quo μ	Optimal
$r_d = 0$	Progressivity (μ)	0.137	0.194	0.137	0.242
	Level (τ)	0.105	0.110	−0.431	−1.035
	Welfare Gain	–	0.19%	0.63%	1.27%
$r_d = (1-\tau_k)r$	Progressivity (μ)	0.137	0.196	0.137	0.239
	Level (τ)	0.105	0.110	−0.400	−0.936
	Welfare Gain	–	0.21%	0.63%	1.03%
$r_d = r$	Progressivity (μ)	0.137	0.197	0.137	0.236
	Level (τ)	0.105	0.110	−0.386	−0.883
	Welfare Gain	–	0.21%	0.64%	0.90%

6.1.2 Basing the LIT on Average Earnings

Under our benchmark implementation of the progressive LIT, households receive large transfers early in life, when their accumulated lifetime earnings are low, and repay these transfers through future taxes as they age and their lifetime earnings increase. We now explore an alternative implementation of the LIT based on average past earnings, which avoids such early-life transfers. We show that our main findings regarding the comparison between AIT and LIT are qualitatively and, to a large extent, quantitatively robust with respect to this alternative implementation. The main drawback of this alternative implementation is that the tax formula becomes more complex and explicitly age-dependent.

We now explain the details of the implementation of a LIT that is based on average earnings. In what follows we assume a zero discount rate for computing lifetime earnings as in the benchmark implementation of the LIT.[47] Let $T(\cdot)$ continue to denote the mapping from pre-tax lifetime earnings to lifetime tax liability. Let Y_{j-1} denote the sum of earnings from age 1 to $j-1$. Average earnings up to age $j-1$ are then defined as $\bar{y}_{j-1} = Y_{j-1}/(j-1)$,

---

[47] As with the benchmark implementation, it is conceptually straightforward to extend this alternative implementation to a nonzero discount rate for computing lifetime income.

and with this definition, average earnings up to age j are given by $\bar{y}_j = Y_j/j = (Y_{j-1} + y_j)/j$.

The alternative implementation of the LIT proceeds in three steps. First, we predict household lifetime earnings over the entire life cycle, assuming that earnings during the remaining working years will remain the same as average earnings up to age j. Second, we compute the predicted lifetime tax liability based on these predicted lifetime earnings, $T(J_R\bar{y}_j)$. Third, we calculate the annualized (over the entire working life) tax liability implied by this predicted lifetime liability as $T(J_R\bar{y}_j)/J_R$, which implies that the accumulated tax liability up to age j is $\frac{j}{J_R}T(J_R\bar{y}_j)$. In this implementation of the LIT households are required to pay this accumulated amount by the end of age j. Therefore, the tax payment due at age j, and assuming that the household has met its tax obligations at age $j-1$, is given by:

$$\widetilde{T}(y, Y, j) = \frac{j}{J_R}T\left(\left[\frac{Y+y}{j}\right]J_R\right) - \frac{j-1}{J_R}T\left(\left[\frac{Y}{j-1}\right]J_R\right).$$

Appendix C.4 shows that in this alternative implementation, lifetime tax liabilities are identical to the benchmark implementation as long as the T function remains the same.

We repeated our main quantitative exercises (in partial equilibrium with the within-cohort budget constraint) under this alternative implementation of the LIT. As Table 7 shows, the welfare gains from switching to a LIT, with status quo- and optimal tax progressivity, respectively, amount to 0.49% and 1.10% of lifetime consumption. These gains are comparable but slightly smaller than in the benchmark implementation (0.63% and 1.27%).[48]

Table 7: Welfare Gains from Annual and Lifetime Income Taxes
(LIT Based on Avg. Earnings)

	Annual Tax		Lifetime Tax	
	Status Quo	Optimal	Status Quo μ	Optimal
A. Benchmark LIT				
Progressivity (μ)	0.137	0.194	0.137	0.242
Level (τ)	0.105	0.110	−0.431	−1.035
Welfare Gain	–	0.19%	0.63%	1.27%
B. LIT Based on Avg. Earnings				
Progressivity (μ)	0.137	0.194	0.137	0.240
Level (τ)	0.105	0.110	−0.477	−1.140
Welfare Gain	–	0.19%	0.49%	1.10%

Therefore, the welfare benefits of lifetime income taxation are largely preserved under this alternative implementation. At the same time, the bottom-left panel of Figure 23 in Appendix C.4 shows that average after-tax earnings now no longer decline sharply over the life cycle under the LIT as the tax liability is now more evenly distributed across ages.

---

[48] The welfare differences between the two implementations arise because Ricardian equivalence does not hold due to the presence of uninsurable idiosyncratic risk, potentially binding borrowing constraints and a gap between the discount rate used to compute lifetime tax liabilities and the private after-tax interest rate.

6.2 Lump-Sum Transfers

The income tax function in our baseline follows the standard HSV form in (2). Recent studies, e.g., Boar and Midrigan (2022), find that adding a lump-sum transfer ι to (2):

$$T(x) = x - (1-\tau)x^{1-\mu} - \iota,$$

provides a better fit to the mapping between pre- and after-tax income, especially at the bottom of the distribution. As a robustness check, we also consider this alternative formulation of the income tax function and compare the performance of AIT and LIT.[49] After estimating the three parameters from CPS data (for details, see Appendix C.5), which yields a lump-sum transfer of $3,920 (in 2016 dollars) together with a reduced progressivity parameter $\mu = 0.093$, we recalibrate the other model parameters under this alternative tax function and repeat the main exercises.[50] As Table 8 shows, the presence of a lump-sum transfer significantly reduces the potential welfare gains from adjusting tax progressivity, consistent with the results of Boar and Midrigan (2022). Even though the findings here are not directly comparable to our benchmark results since the status quo policy has been re-estimated, our qualitative conclusion nevertheless remains unchanged: switching from an AIT to a LIT, whether with the status quo or optimal progressivity, yields substantial welfare gains.

Table 8: Welfare Gains from AIT and LIT (With Lump-Sum Transfers)

	Annual Tax		Lifetime Tax	
	Status Quo	Optimal	Status Quo μ	Optimal
A. Baseline HSV Tax Functions				
Progressivity (μ)	0.137	0.194	0.137	0.242
Level (τ)	0.105	0.110	−0.431	−1.035
Welfare Gain	–	0.19%	0.63%	1.27%
B. HSV Tax Functions with Lump-Sum Transfers				
Progressivity (μ)	0.093	0.099	0.093	0.150
Level (τ)	0.196	0.196	−0.104	−0.341
Lump-Sum (ι)	0.072	0.072	0.072	0.072
Welfare Gain	–	0.00%	0.40%	0.55%

Notes: For panel B, model parameters are recalibrated targeting the same empirical moments as the baseline model.

6.3 Accounting for General Equilibrium Effects

Our baseline analysis assumes that the interest rate r and the wage rate w are not affected by tax reforms. We now extend the analysis to general equilibrium where (r, w) are determined by domestic capital and labor market clearing and thus are endogenous to tax

---

[49] For both AIT and LIT, ι represents the lump-sum transfers per year.

[50] To keep the computational burden manageable, we do not adjust the lump-sum component in tax reforms, only the progressivity and level parameters.

policy.[51] Table 9 shows that accounting for general equilibrium leads to substantially more progressive optimal policies.[52] As in our baseline results, the optimal LIT is still slightly more progressive than the optimal AIT (Panel A). For both systems, general equilibrium effects raise wages and lower interest rates, with magnitudes of changes that are close for the two types of tax systems (Panel B).

Table 9: Optimal Income Tax in General Equilibrium

	Annual Tax	Lifetime Tax		Annual Tax	Lifetime Tax
A. Optimal Income Tax			*B. General Equilibrium Effects*		
Progressivity (μ)	0.345	0.371	Wage	5.63%	5.68%
Level (τ)	0.086	-2.415	Interest Rate^{a}	$-1.47pp$	$-1.48pp$
Level (comparable)	8.6%	14.6%			
Avg. Tax Rate	14.6%	18.3%			
C. Aggregate and Distributional Implications			*D. Welfare Gains and Decomposition*		
Hours	−7.38%	−10.19%	Total Welfare Gain	5.84%	7.39%
Earnings	−2.87%	−2.86%	Consumption	−0.81%	−0.27%
Consumption	−8.30%	−8.25%	Level	−9.08%	−9.12%
			Age	2.77%	2.96%
Labor Efficiency	4.86%	8.16%	Distribution	5.50%	5.89%
			Labor	6.65%	7.67%
Hours Inequality	−11.45%	8.90%	Level	5.57%	7.72%
Earnings Inequality	−5.56%	12.22%	Age	0.41%	0.36%
Consumption Inequality	−39.28%	−40.51%	Distribution	0.67%	−0.41%

Notes: ^{a} The number reported is the percentage-point change in interest rate.

The optimal income taxes still reduce total hours and earnings, but compared to our baseline results, total earnings fall less due to the higher wage level in general equilibrium (Panel C). For the same reason, aggregate labor efficiency improves under both the annual and lifetime income taxes. However, labor efficiency grows less than the wage under the AIT, but more under the LIT. Therefore, consistent with our previous findings, the optimal annual (lifetime) income tax still induces a shift in the distribution of hours that decreases (increases) aggregate efficiency, and hours- and earnings inequality. Because the optimal policies in general equilibrium are more progressive, consumption inequality falls more than in our baseline results. The welfare gains from the optimal annual and lifetime income taxes are substantially larger in general equilibrium, but as in our baseline analysis, the LIT still outperforms the AIT by about 1.55% in terms of welfare gains (Panel D). And most of this extra welfare gain from the optimal LIT is again due to a greater increase in leisure.

Appendix C.6.2 further shows that the larger welfare gains in Panel D of Table 9 mostly stem from the direct and indirect effects of factor price adjustments, while the welfare advantage of the optimal LIT relative to the optimal AIT is mainly driven by the direct effects

---

[51] Since we impose a within-cohort government budget constraint when solving for optimal policy, the level of government debt must adjust to balance the government period budget in stationary equilibrium.

[52] Appendix C.6.1 discusses the life-cycle implications of tax reforms with general equilibrium effects.

of tax policy already present in partial equilibrium, justifying our choice of this specification as our benchmark.

7 Conclusion

In this paper, we have explored whether an income tax system based on household lifetime earnings is superior to the prevailing annual income tax system adopted by most countries. This inquiry arises from the fundamental observation that lifetime earnings provide a potentially more precise indicator of household welfare, making it a more suitable base for welfare redistribution. Our quantitative analysis, based on an incomplete-markets life-cycle model of heterogeneous households calibrated to the U.S. economy, yields an affirmative answer. Switching to lifetime income tax promotes aggregate labor efficiency by inducing a more efficient distribution of hours worked. A more progressive lifetime income tax further enhances this benefit, whereas a more progressive annual income tax has the opposite effect. Consequently, the optimal lifetime income tax is more progressive and achieves larger welfare gains than the optimal annual income tax.

To arrive at this conclusion we made several important simplifying assumptions that deserve further scrutiny. First, the idiosyncratic labor productivity process households face is exogenous and policy invariant, abstracting from endogenous schooling and human capital accumulation decisions. Human capital investments often involve substantial income variation over the life cycle, e.g., pursuing a Ph.D. with little or no income initially in exchange for higher earnings later. We conjecture that a LIT is even more advantageous in the presence of such decisions, relative to an AIT, since the latter tends to penalize year-to-year income variation through higher average tax rates over the lifetime, as we have shown above.

Second, we abstracted from endogenous family formation and intra-household allocation decisions that could potentially be affected by the tax system. If the tax unit is the household, the benefits of forming a dual-earner household could be reduced by a LIT since with such a tax system each individual can effectively pool earnings across the own life cycle for tax purposes, and income pooling across earners becomes less important. For the same reason, in response to transitory earnings shocks of one member the labor supply response of the other member, i.e., the added-worker effect, would likely be smaller under a LIT.

Finally, we have abstracted from the endogenous intergenerational transmission of wealth between generations, and the associated taxation of bequests, and wealth more generally.[53] It is an open question how a reform from an AIT towards a lifetime tax on earnings interacts with the taxation of capital. We hope our paper helps to reignite discussion on lifetime income taxation and stimulate future research along these dimensions.

---

[53] See, e.g., Guvenen, Kambourov, Kuruscu, Ocampo, and Chen (2023) for a recent study of this issue.

References

Ábrahám, Árpád, Pavel Brendler, and Eva Cárceles-Poveda (2024). “Optimal Income Redistribution”. *Working paper.*

Aiyagari, S. Rao (1994). “Uninsured Idiosyncratic Risk and Aggregate Saving”. *The Quarterly Journal of Economics* 109.3, pp. 659–684.

Albanesi, Stefania and Christopher Sleet (2006). “Dynamic Optimal Taxation with Private Information”. *The Review of Economic Studies* 73.1, pp. 1–30.

Arellano, Manuel, Richard Blundell, and Stéphane Bonhomme (2017). “Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework”. *Econometrica* 85.3, pp. 693–734.

Audoly, Richard, Rory McGee, Sergio Ocampo-Diaz, and Gonzalo Paz-Pardo (2025). “A Practitioner’s Note on the Shapley-Owen-Shorrocks Decomposition”. Working Paper.

Auerbach, Alan J., Laurence J. Kotlikoff, and Darryl Koehler (2023). “US Inequality and Fiscal Progressivity: An Intragenerational Accounting”. *Journal of Political Economy* 131.5, pp. 1249–1293.

Battaglini, Marco and Stephen Coate (2008). “Pareto efficient income taxation with stochastic abilities”. *Journal of Public Economics* 92.3, pp. 844–868.

Batzer, Ross (2022). “Essays on Income Taxation”. Ph.D. dissertation. University of Minnesota.

Bewley, Truman (1986). “Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers”. *Contributions to Mathematical Economics in Honor of Gerard Debreu*, pp. 79–102.

Bick, Alexander, Adam Blandin, and Richard Rogerson (2024). “Hours Worked and Lifetime Earnings Inequality”. *Working paper.*

Blomquist, Sören and Luca Micheletto (2008). “Age-Related Optimal Income Taxation”. *The Scandinavian Journal of Economics* 110.1, pp. 45–71.

Blundell, Richard, Luigi Pistaferri, and Ian Preston (2008). “Consumption Inequality and Partial Insurance”. *American Economic Review* 98.5, 1887–1921.

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten (2016). “Consumption Inequality and Family Labor Supply”. *American Economic Review* 106.2, 387–435.

Boar, Corina and Virgiliu Midrigan (2022). “Efficient redistribution”. *Journal of Monetary Economics* 131, pp. 78–91.

Bénabou, Roland (2002). “Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?” *Econometrica* 70.2, pp. 481–517.

Cagetti, Marco and Mariacristina De Nardi (2006). “Entrepreneurship, Frictions, and Wealth”. *Journal of Political Economy* 114.5, pp. 835–870.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger (2009). “Taxing Capital? Not a Bad Idea after All!” *American Economic Review* 99.1, pp. 25–48.

Conesa, Juan Carlos and Dirk Krueger (2006). “On the Optimal Progressivity of the Income Tax Code”. *Journal of Monetary Economics* 53.7, pp. 1425–1450.

De Nardi, Mariacristina, Giulio Fella, and Gonzalo Paz-Pardo (2019). “Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare”. *Journal of the European Economic Association* 18.2, pp. 890–926.

Dyrda, Sebastian and Marcelo Pedroni (2022). “Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk”. *The Review of Economic Studies* 90.2, pp. 744–780.

Dávila, Julio, Jay H. Hong, Per Krusell, and José-Víctor Ríos-Rull (2012). “Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks”. *Econometrica* 80.6, pp. 2431–2467.

Erosa, Andrés and Martin Gervais (2002). “Optimal Taxation in Life-Cycle Economies”. *Journal of Economic Theory* 105.2, pp. 338–369.

Farhi, Emmanuel and Iván Werning (2007). “Inequality and Social Discounting”. *Journal of Political Economy* 115.3, pp. 365–402.

— (2013). “Insurance and Taxation over the Life Cycle”. *The Review of Economic Studies* 80.2, pp. 596–635.

Fehr, Hans, Manuel Kallweit, and Fabian Kindermann (2013). “Should pensions be progressive?” *European Economic Review* 63, pp. 94–116.

Garriga, Carlos (2001). “Optimal Fiscal Policy in Overlapping Generations Models”. *Working paper.*

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski (2003). “Optimal Indirect and Capital Taxation”. *The Review of Economic Studies* 70.3, pp. 569–587.

Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski (2016). “Redistribution and Social Insurance”. *American Economic Review* 106.2, 359–86.

Grochulski, Borys and Narayana Kocherlakota (2010). “Nonseparable preferences and optimal social security systems”. *Journal of Economic Theory* 145.6, pp. 2055–2077.

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen (2023). “Use It or Lose It: Efficiency and Redistributional Effects of Wealth Taxation”. *The Quarterly Journal of Economics* 138.2, pp. 835–894.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song (2021). “What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics?” *Econometrica* 89.5, pp. 2303–2339.

Guvenen, Fatih, Serdar Ozkan, and Rocio Madera (2024). “Consumption dynamics and welfare under non-Gaussian earnings risk”. *Journal of Economic Dynamics and Control* 169. Heterogeneity and Macroeconomics of Labor Markets, p. 104945.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante (2005). “Two Views of Inequality over the Life Cycle”. *Journal of the European Economic Association* 3.2-3, pp. 765–775.

— (2014). “Consumption and Labor Supply with Partial Insurance: An Analytical Framework”. *American Economic Review* 104.7, pp. 2075–2126.

— (2017). “Optimal Tax Progressivity: An Analytical Framework”. *The Quarterly Journal of Economics* 132.4, pp. 1693–1754.

— (2020). “Optimal progressivity with age-dependent taxation”. *Journal of Public Economics* 189, p. 104074.

Holter, Hans A., Dirk Krueger, and Serhiy Stepanchuk (2023). “Until the IRS Do Us Part: (Optimal) Taxation of Households”. *Working paper.*

Huggett, Mark (1993). “The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies”. *Journal of Economic Dynamics and Control* 17.5-6, pp. 953 –969.

Huggett, Mark and Greg Kaplan (2011). “Human capital values and returns: Bounds implied by earnings and asset returns data”. *Journal of Economic Theory* 146.3, pp. 897–919.

Huggett, Mark and Juan Carlos Parra (2010). “How Well Does the U.S. Social Insurance System Provide Social Insurance?” *Journal of Political Economy* 118.1, pp. 76–112.

Kapička, Marek (2020). “Quantifying the Welfare Gains from History Dependent Income Taxation”. *Working paper.*

Kaplan, Greg (2012). “Inequality and the life cycle”. *Quantitative Economics* 3.3, pp. 471–525.

Kaplan, Greg and Giovanni L. Violante (2010). “How Much Consumption Insurance beyond Self-Insurance?” *American Economic Journal: Macroeconomics* 2.4, 53–87.

Kirkby, Robert (2025). “Discretizing earnings dynamics: implications of Gaussian-mixture shocks for life-cycle models”. *Japanese Economic Review* 76, pp. 493–519.

Kocherlakota, Narayana R. (2005). “Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation”. *Econometrica* 73.5, pp. 1587–1621.

Kremer, Michael (2001). “Should Taxes be Independent of Age?” *Working paper.*

Mendoza, Enrique G., Assaf Razin, and Linda L. Tesar (1994). “Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption”. *Journal of Monetary Economics* 34.3, pp. 297 –323.

Moschini, Emily G. and Monica Tran-Xuan (2025). “Family policies and child skill accumulation”. *Review of Economic Dynamics* 56, p. 101270.

Quadrini, Vincenzo (1999). “The Importance of Entrepreneurship for Wealth Concentration and Mobility”. *Review of Income and Wealth* 45.1, pp. 1–19.

— (2000). “Entrepreneurship, Saving, and Social Mobility”. *Review of Economic Dynamics* 3.1, pp. 1–40.

Trabandt, Mathias and Harald Uhlig (2011). “The Laffer curve revisited”. *Journal of Monetary Economics* 58.4, pp. 305 –327.

Urquizo, César (2025). “Lifetime Hours Inequality and Occupational Choice”. *Working paper.*

Vickrey, William (1939). “Averaging of Income for Income-Tax Purposes”. *Journal of Political Economy* 47.3, pp. 379–397.

Weinzierl, Matthew (2011). “The Surprising Power of Age-Dependent Taxes”. *The Review of Economic Studies* 78.4, pp. 1490–1518.

Werning, Iván (2007). “Optimal Fiscal Policy with Redistribution*”. *The Quarterly Journal of Economics* 122.3, pp. 925–967.

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

Online Appendix

Taxes on Lifetime Income: A Good Idea?

Dirk Krueger and Chunzan Wu

A Proofs for the Two-Period Model

A.1 Proof of Proposition 1

The Lagrange function for the social planner's problem is

$$\mathcal{L} = U(c_1, l_1) + \mathbf{E}\{U(c_2, l_2)\} - \lambda[c_1 + \mathbf{E}\{c_2\} - w_1 l_1 - \mathbf{E}\{w_2 l_2\}],$$

where λ is the multiplier on the resource constraint. Note that c_2 and l_2 are all contingent on the second period state. The first order conditions are then

$$\begin{aligned} c_1 : &\quad U_c(c_1, l_1) = \lambda, \\ c_2(s) : &\quad U_c(c_2(s), l_2(s)) = \lambda, \\ l_1 : &\quad U_l(c_1, l_1) = -\lambda w_1, \\ l_2(s) : &\quad U_l(c_2(s), l_2(s)) = -\lambda w_2(s), \end{aligned}$$

together with the resource constraint (3).

Combining the first order conditions w.r.t. c_1 and $c_2(s)$ to eliminate λ, we get the Euler equation (4). Using the first order conditions w.r.t. c_1 and $c_2(s)$ to substitute λ in the first order conditions w.r.t. l_1 and $l_2(s)$, respectively, we have the intratemporal optimality conditions (5). Finally, combining the first order conditions w.r.t. l_1 and $l_2(s)$ to eliminate λ, we get (6) and (7).

A.1.1 Proof of Corollary 1

Follows directly by taking the limit $\sigma \to 0$ in equation (8).

A.2 Proof of Proposition 2

Note that in partial equilibrium with exogenous wage and interest rate and no government, the only equilibrium conditions are those associated with the household utility maximization. The Lagrange function for the household's problem is

$$\mathcal{L} = U(c_1, l_1) + \mathbf{E}\{U(c_2(s), l_2(s))\} - \lambda_1(c_1 + a - w_1 l_1) - \mathbf{E}\{\lambda_2(s)[c_2(s) - a - w_2(s) l_2(s)]\},$$

where λ_1 and $\lambda_2(s)$ are multipliers on the first and second period budget constraints. The first order conditions are then

$$c_1 : \quad U_c(c_1, l_1) = \lambda_1,$$

$$
\begin{aligned}
c_2(s):&\quad U_c(c_2(s),l_2(s))=\lambda_2(s),\\
l_1:&\quad U_l(c_1,l_1)=-\lambda_1 w_1,\\
l_2(s):&\quad U_l(c_2(s),l_2(s))=-\lambda_2(s)w_2(s),\\
a:&\quad \lambda_1=\mathbf{E}\{\lambda_2(s)\},
\end{aligned}
$$

together with the first and second period budget constraints.

Use the conditions w.r.t. c_1 and $c_2(s)$ to substitute λ_1 and $\lambda_2(s)$ in the condition w.r.t. a, and we get the Euler equation (9). Use the first order conditions w.r.t. c_1 and $c_2(s)$ to substitute λ_1 and $\lambda_2(s)$ in the first order conditions w.r.t. l_1 and $l_2(s)$, respectively, and we have the intratemporal optimality conditions (10). Use the conditions w.r.t. l_1 and $l_2(s)$ to substitute λ_1 and $\lambda_2(s)$ in the condition w.r.t. a, and we get (11). Use the the condition w.r.t $c_2(s)$ to substitute $\lambda_2(s)$ in the condition w.r.t. $l_2(s)$, and then take the ratio of the resulting equation between any two states s' and s, and we get (12).

A.2.1 Proof of Corollary 2

Follows directly from equation (10), using the fact that with $\sigma=0$ we have $U_c=1$, and comparing the resulting expression for labor supply with that in Corollary 1.

A.3 Proof of Proposition 3

The Ramsey equilibrium consists of the optimality conditions of households and the government budget constraint. The Lagrange function for the household's optimization problem is

$$
\mathcal{L}=U(c_1,l_1)+\mathbf{E}\{U(c_2,l_2)\}-\lambda_1[c_1+a-w_1l_1+T(w_1l_1)]-\mathbf{E}\{\lambda_2[c_2-a-w_2l_2+\widetilde{T}(w_1l_1,w_2l_1)]\},
$$

where λ_1 and λ_2 are multipliers on the first and second period budget constraints. Note that c_2, l_2, and λ_2 are all contingent on the second period state s, whereas a is state-uncontingent. The first order conditions are then

$$
\begin{aligned}
c_1:&\quad U_c(c_1,l_1)=\lambda_1, &(21)\\
c_2:&\quad U_c(c_2,l_2)=\lambda_2, &(22)\\
l_1:&\quad -U_l(c_1,l_1)=\lambda_1 w_1[1-T'(w_1l_1)]-\mathbf{E}\{\lambda_2 w_1\widetilde{T}_1(w_1l_1,w_2l_2)\}, &(23)\\
l_2:&\quad -U_l(c_2,l_2)=\lambda_2 w_2[1-\widetilde{T}_2(w_1l_1,w_2l_2)], &(24)\\
a:&\quad \lambda_1=\mathbf{E}\{\lambda_2\}, &(25)
\end{aligned}
$$

together with the first and second period budget constraints.

A.3.1 Annual Income Tax (AIT)

Under annual income tax,

$$
\widetilde{T}(w_1l_1,w_2l_2)=T(w_2l_2).
$$

⇒

$$\begin{aligned}\widetilde{T}_1(w_1l_1, w_2l_2) &= 0,\\ \widetilde{T}_2(w_1l_1, w_2l_2) &= T'(w_2l_2).\end{aligned}$$

Use (23) and (24) to eliminate λ_1 and λ_2 in (25), and we have the intertemporal condition:

$$\frac{U_l(c_1,l_1)}{w_1[1-T'(w_1l_1)]} = \mathbf{E}\left\{\frac{U_l(c_2,l_2)}{w_2[1-T'(w_2l_2)]}\right\}.$$

Use (22) to eliminate λ_2 in (24), then take the ratio across two possible second period states s' and s, and we have

$$\frac{U_l(c_2(s'),l_2(s'))}{U_l(c_2(s),l_2(s))} = \left[\frac{1-T'(w_2(s')l_2(s'))}{1-T'(w_2(s)l_2(s))}\right]\left[\frac{U_c(c_2(s'),l_2(s'))}{U_c(c_2(s),l_2(s))}\right]\left[\frac{w_2(s')}{w_2(s)}\right].$$

A.3.2 Lifetime Income Tax (LIT)

Under lifetime income tax,

$$\widetilde{T}(w_1l_1, w_2l_2) = T(w_1l_1 + w_2l_2) - T(w_1l_1).$$

⇒

$$\begin{aligned}\widetilde{T}_1(w_1l_1, w_2l_2) &= T'(w_1l_1 + w_2l_2) - T'(w_1l_1),\\ \widetilde{T}_2(w_1l_1, w_2l_2) &= T'(w_1l_1 + w_2l_2).\end{aligned}$$

(23) then becomes

$$\begin{aligned}-U_l(c_1,l_1) &= \lambda_1 w_1[1-T'(w_1l_1)] - \mathbf{E}\{\lambda_2 w_1[T'(w_1l_1+w_2l_2) - T'(w_1l_1)]\}\\ &= \lambda_1 w_1[1-T'(w_1l_1)] - \mathbf{E}\{\lambda_2 w_1[-1+T'(w_1l_1+w_2l_2) + 1 - T'(w_1l_1)]\}\\ &= w_1(\lambda_1 - \mathbf{E}\{\lambda_2\})[1-T'(w_1l_1)] + w_1\mathbf{E}\{\lambda_2[1-T'(w_1l_1+w_2l_2)]\})\end{aligned}$$

Use (25) to eliminate $\lambda_1 - \mathbf{E}\{\lambda_2\}$, and (24) to eliminate $\lambda_2[1-T'(w_1l_1+w_2l_2)]$, then divide both sides by $-w_1$, and we have the intertemporal condition:

$$\frac{U_l(c_1,l_1)}{w_1} = \mathbf{E}\left\{\frac{U_l(c_2,l_2)}{w_2}\right\}.$$

Use (22) to eliminate λ_2 in (24), then take the ratio across two possible second period states s' and s, and we have

$$\frac{U_l(c_2(s'),l_2(s'))}{U_l(c_2(s),l_2(s))} = \left[\frac{1-T'(w_1l_1+w_2(s')l_2(s'))}{1-T'(w_1l_1+w_2(s)l_2(s))}\right]\left[\frac{U_c(c_2(s'),l_2(s'))}{U_c(c_2(s),l_2(s))}\right]\left[\frac{w_2(s')}{w_2(s)}\right].$$

A.3.3 Proof of Corollary 3

The statement that the intertemporal allocation of labor is efficient under the LIT follows directly from Proposition 3 (and Proposition 1).

In the absence of wage risk, and with additively separable utility and HSV tax functions, the optimality conditions under the AIT are:

$$\begin{aligned} {c_1}^{-\sigma} &= {c_2}^{-\sigma}, \\ {c_1}^{-\sigma}(1-\tau)(1-\mu)(w_1 l_1)^{-\mu} w_1 &= \psi l_1^{1/\eta}, \\ c_2^{-\sigma}(1-\tau)(1-\mu)(w_2 l_2)^{-\mu} w_2 &= \psi l_2^{1/\eta}. \end{aligned}$$

Hence,

$$\frac{l_2}{l_1} = \left(\frac{w_2}{w_1}\right)^{\eta \frac{1-\mu}{1+\mu\eta}}.$$

Aggregate labor efficiency under the AIT is then given by:

$$ALE^{AIT} \equiv \frac{w_1 l_1 + w_2 l_2}{l_1 + l_2} = \frac{w_1 + w_2 \left(\frac{w_2}{w_1}\right)^{\eta \frac{1-\mu}{1+\mu\eta}}}{1 + \left(\frac{w_2}{w_1}\right)^{\eta \frac{1-\mu}{1+\mu\eta}}} = \frac{w_1 - w_2}{1 + \left(\frac{w_2}{w_1}\right)^{\eta \frac{1-\mu}{1+\mu\eta}}} + w_2.$$

With a LIT, the optimality conditions are:

$$\begin{aligned} {c_1}^{-\sigma} &= {c_2}^{-\sigma}, \\ {c_1}^{-\sigma}(1-\tau)(1-\mu)(w_1 l_1 + w_2 l_2)^{-\mu} w_1 &= \psi l_1^{1/\eta}, \\ c_2^{-\sigma}(1-\tau)(1-\mu)(w_1 l_1 + w_2 l_2)^{-\mu} w_2 &= \psi l_2^{1/\eta}. \end{aligned}$$

Hence,

$$\frac{l_2}{l_1} = \left(\frac{w_2}{w_1}\right)^{\eta}.$$

Therefore, aggregate labor efficiency under the LIT is given by:

$$ALE^{LIT} \equiv \frac{w_1 l_1 + w_2 l_2}{l_1 + l_2} = \frac{w_1 + w_2 \left(\frac{w_2}{w_1}\right)^{\eta}}{1 + \left(\frac{w_2}{w_1}\right)^{\eta}} = \frac{w_1 - w_2}{1 + \left(\frac{w_2}{w_1}\right)^{\eta}} + w_2.$$

If $\mu = 0$, obviously $ALE^{LIT} = ALE^{AIT}$. If $\mu > 0$, then:

1. When $w_1 = w_2$, $ALE^{LIT} = ALE^{AIT} = w_2$;
2. When $w_1 > w_2$, $\frac{w_2}{w_1} < 1$, and thus $\left(\frac{w_2}{w_1}\right)^{\eta \frac{1-\mu}{1+\mu\eta}} > \left(\frac{w_2}{w_1}\right)^{\eta}$; since $w_1 - w_2 > 0$, we have $ALE^{LIT} > ALE^{AIT} > w_2$;
3. When $w_1 < w_2$, $\frac{w_2}{w_1} > 1$, and thus $\left(\frac{w_2}{w_1}\right)^{\eta \frac{1-\mu}{1+\mu\eta}} < \left(\frac{w_2}{w_1}\right)^{\eta}$; since $w_1 - w_2 < 0$, we have $w_2 > ALE^{LIT} > ALE^{AIT}$;

which proves the corollary.

A.3.4 Proof of Corollary 4

The results in the corollary for $\sigma = 0$ and the tax function $T(y) = y - (1-\tau)y^{1-\mu}$ directly follow from Proposition 3 for the AIT. For the LIT, the household first order conditions imply that

$$\psi l_1^{1/\eta} = \mathbf{E}\left\{(1-\tau)(1-\mu)[w_1 l_1 + w_2(\tilde{s}) l_2(\tilde{s})]^{-\mu}\right\} w_1, \tag{26}$$

and

$$\psi [l_2(s)]^{1/\eta} = (1-\tau)(1-\mu)[w_1 l_1 + w_2(s) l_2(s)]^{-\mu} w_2(s),$$

where the expectation is taken with respect to the random state $\tilde{s}$, and s refers to a specific state in the second period. Taking the ratio of the two equations, we have

$$\left(\frac{l_2(s)}{l_1}\right)^{1/\eta} = \left(\frac{w_2(s)}{w_1}\right) \frac{[w_1 l_1 + w_2(s) l_2(s)]^{-\mu}}{\mathbf{E}\{[w_1 l_1 + w_2(\tilde{s}) l_2(\tilde{s})]^{-\mu}\}}.$$

Since there is no uncertainty in the first period, we can divide both the numerator and the denominator of the right-hand side by $(w_1 l_1)^{-\mu}$ to rewrite this expression as

$$\left(\frac{l_2(s)}{l_1}\right)^{1/\eta} = \left(\frac{w_2(s)}{w_1}\right) \frac{[1 + \frac{w_2(s)}{w_1}\frac{l_2(s)}{l_1}]^{-\mu}}{\mathbf{E}\{[1 + \frac{w_2(\tilde{s})}{w_1}\frac{l_2(\tilde{s})}{l_1}]^{-\mu}\}}.$$

Note that the expectation $\mathbf{E}\{\cdot\}$ on the right-hand side is a constant as we vary $\ln(w_2(s')/w_1)$ and $\ln(w_2(s')/w_2(s))$ since the distribution of $\tilde{s}$ is atomless and thus the expectation is unaffected by changes in l_2 or w_2 for any specific realization of $\tilde{s}$, and from (26) labor l_1 in the first period does not change (since by assumption w_1 does not change, see footnote 14).

For simplicity, define $y = l_2(s)/l_1$, and $x = w_2(s)/w_1$, and the equation can be written as

$$y^{1/\eta} = x \frac{(1+xy)^{-\mu}}{\texttt{CONSTANT}},$$

$\Rightarrow$

$$\frac{1}{\eta} d\ln(y) = d\ln(x) - \mu \frac{xy}{1+xy}[d\ln(x) + d\ ln(y)],$$

$\Rightarrow$

$$\frac{d\ln(y)}{d\ln(x)} = \eta \left(\frac{1 - \mu\frac{xy}{1+xy}}{1 + \mu\eta\frac{xy}{1+xy}}\right).$$

Since $xy/(1+xy) \in (0,1)$, we have

$$\underbrace{\eta\left(\frac{1-\mu}{1+\mu\eta}\right)}_{\text{AIT}} < \underbrace{\frac{d\ln(y)}{d\ln(x)}}_{\text{LIT}} < \underbrace{\eta}_{\text{Efficient Alloc.}},$$

and the elasticity of *relative* labor supply $l_2(s)/l_1$ with respect to the *relative* wage across the two periods $w_2(s)/w_1$ under the LIT is smaller than under the efficient allocation, but closer to it than under the AIT, as stated in the corollary. And if in addition S is a singleton (no risk in the second period), then the relative labor allocation is fully efficient under the LIT:

$$\frac{l_2}{l_1} = \left(\frac{w_2}{w_1}\right)^{\eta},$$

and the elasticity of relative labor to relative wages is η.

Now consider the effect of $d\ln\left(\frac{w_2(s')}{w_2(s)}\right)$ on the labor allocation across two states s, s' in the second period, driven by a change in $w_2(s')$. The first-order conditions imply that

$$\left(\frac{l_2(s')}{l_2(s)}\right)^{1/\eta} = \frac{w_2(s')}{w_2(s)}\left[\frac{w_1 l_1 + w_2(s')l_2(s')}{w_1 l_1 + w_2(s)l_2(s)}\right]^{-\mu}.$$

Taking logs of both sides and then the total differential, and recognizing that by assumption $d\ln(w_2(s)) = 0$ and thus, from the first order conditions, $d\ln(l_1) = d\ln(l_2(s)) = 0$,

$$\frac{1}{\eta}d\ln\left(\frac{l_2(s')}{l_2(s)}\right) = d\ln\left(\frac{w_2(s')}{w_2(s)}\right) - \mu\frac{w_2(s')l_2(s')}{w_1 l_1 + w_2(s')l_2(s')}[dln(w_2(s')) + dln(l_2(s'))].$$

Therefore,

$$\varepsilon_{s',s} = \frac{d\ln\left(\frac{l_2(s')}{l_2(s)}\right)}{d\ln\left(\frac{w_2(s')}{w_2(s)}\right)} = \frac{1-\mu\frac{w_2(s')l_2(s')}{w_1l_1+w_2(s')l_2(s')}}{\frac{1}{\eta}+\mu\frac{w_2(s')l_2(s')}{w_1l_1+w_2(s')l_2(s')}} \in \left(\eta\frac{1-\mu}{1+\mu\eta}, \eta\right),$$

and thus relative labor is more elastic to relative wages under the LIT than under the AIT (and closer to the efficient allocation of labor).

Note that the result does not generalize to arbitrary changes in $w_2(s')$ and $w_2(s)$. The main difficulty is that multiple combinations of changes in $w_2(s')$ and $w_2(s)$ produce the same change in the ratio $w_2(s')/w_2(s)$ but lead to different changes in $l_2(s')/l_2(s)$, since under a LIT, the marginal tax rate depends on $w_1 l_1 + w_2(\tilde{s})l_2(\tilde{s})$.

A.4 Effect of τ: Additively Separable Preferences and HSV Tax Function

We now show that under the additional assumptions about the two-period model in Section B, a change in the tax level parameter τ alone does not affect the relative distribution of consumption or labor over time and across states for either annual or lifetime income taxes, but only shifts the overall levels of consumption and labor. As a result, all wedges defined in Table 10, except the consumption-labor wedge, remain unaffected.

Proposition 4. *Holding the income tax type fixed (either annual or lifetime income tax), let $\{\bar{c}_1, \bar{l}_1, [\bar{c}_2(s), \bar{l}_2(s)]_{s\in S}\}$ denote the household's optimal choices under a tax policy with level $\tau = 0$ and progressivity μ. Then under a tax policy characterized by τ and μ, the household's optimal choices satisfy*

$$\frac{c_1}{\bar{c}_1} = \frac{c_2(s)}{\bar{c}_2(s)} = \kappa_c, \ \forall s \in S,$$
$$\frac{l_1}{\bar{l}_1} = \frac{l_2(s)}{\bar{l}_2(s)} = \kappa_l, \ \forall s \in S,$$

where

$$\kappa_c = (1-\tau)^{1+\frac{(1-\sigma)(1-\mu)}{\mu+1/\eta+(1-\mu)\sigma}},$$
$$\kappa_l = (1-\tau)^{\frac{1-\sigma}{\mu+1/\eta+(1-\mu)\sigma}}.$$

Proof.

Annual income tax. The optimal choices under (τ, μ) are characterized by

$$
\begin{aligned}
{c_1}^{-\sigma} &= \mathbf{E}\{(c_2(s))^{-\sigma}\}, \\
{c_1}^{-\sigma}(1-\tau)(1-\mu)(w_1 l_1)^{-\mu} w_1 &= \psi l_1^{1/\eta}, \\
[c_2(s)]^{-\sigma}(1-\tau)(1-\mu)[w_2(s) l_2(s)]^{-\mu} w_2(s) &= \psi [l_2(s)]^{1/\eta}, \ \forall s \in S, \\
c_1 + c_2(s) &= (1-\tau)(w_1 l_1)^{1-\mu} + (1-\tau)[w_2(s) l_2(s)]^{1-\mu}, \ \forall s \in S.
\end{aligned}
$$

Plug in the proposed solution, and we have

$$
\begin{aligned}
\kappa_c^{-\sigma}(1-\tau) &= \kappa_l^{\mu+1/\eta}, \\
\kappa_c &= (1-\tau)\kappa_l^{1-\mu}.
\end{aligned}
$$

$\Rightarrow$

$$
\begin{aligned}
\kappa_c =& (1-\tau)^{1+\frac{(1-\sigma)(1-\mu)}{\mu+1/\eta+(1-\mu)\sigma}}, \\
\kappa_l =& (1-\tau)^{\frac{1-\sigma}{\mu+1/\eta+(1-\mu)\sigma}}.
\end{aligned}
$$

Lifetime income tax. The optimal choices under (τ, μ) are characterized by

$$
\begin{aligned}
{c_1}^{-\sigma} &= \mathbf{E}\{(c_2(s))^{-\sigma}\}, \\
\mathbf{E}\left\{[c_2(s)]^{-\sigma}(1-\tau)(1-\mu)[w_1 l_1 + w_2(s) l_2(s)]^{-\mu}\right\} w_1 &= \psi l_1^{1/\eta}, \\
[c_2(s)]^{-\sigma}(1-\tau)(1-\mu)[w_1 l_1 + w_2(s) l_2(s)]^{-\mu} w_2(s) &= \psi [l_2(s)]^{1/\eta}, \ \forall s \in S, \\
c_1 + c_2(s) &= (1-\tau)[w_1 l_1 + w_2(s) l_2(s)]^{1-\mu}, \ \forall s \in S.
\end{aligned}
$$

Plug in the proposed solution, and we have

$$
\begin{aligned}
\kappa_c^{-\sigma}(1-\tau) &= \kappa_l^{\mu+1/\eta}, \\
\kappa_c &= (1-\tau)\kappa_l^{1-\mu}.
\end{aligned}
$$

Note that these conditions are the same as those for annual income tax. Therefore, κ_c and κ_l are the same for both annual and lifetime income taxes.

□

Corollary 5. *For both annual and lifetime income taxes, the intertemporal and dispersion wedges of consumption and labor remain invariant with respect to the tax level parameter* τ*. Only the consumption-labor wedge and the overall levels of consumption and labor are affected.*

B Annual vs. Lifetime Income Taxes: Quantitative Exploration of the Simple Model

The efficient allocation in Section 2.2 features: i) an efficient distribution of labor, with equal marginal disutility of earnings U_l/w over the life cycle and across idiosyncratic states; ii) an efficient distribution of consumption, with equal marginal utility of consumption U_c over the life cycle and across idiosyncratic states; and iii) efficient levels of consumption and labor such that the marginal disutility of earnings equals the marginal utility of consumption.

Due to incomplete markets and distortionary taxation, the market equilibrium in laissez-faire or under annual- or lifetime income taxes fails to attain the efficient allocation, resulting in lower welfare levels than with the efficient allocation. To quantify these deviations and the severity of distortions causing them, we introduce the following five wedges. For transparency of the exposition of the results, we further assume that the second-period wage follows an equal-probability, two-point distribution with values $w_2^H > w_2^L$.

1. *Intertemporal labor wedge*: a "tax" on relocating labor (earnings) from first to second period. For each unit reduction in first-period earnings, the required expected earnings increase in the second period to compensate the first-period earnings loss is $1 + \chi_{\text{lab}}^{\text{LC}}$:

$$\frac{U_{l_1}}{w_1} = (1 + \chi_{\text{lab}}^{\text{LC}})\mathbf{E}\left\{\frac{U_{l_2}}{w_2}\right\}.$$

2. *Labor dispersion wedge*: a "tax" on relocating labor (earnings) from the low- to high-wage state in the second period. For each unit of earnings reduction in the low-wage state, the required earnings increase in the high-wage state is $1 + \chi_{\text{lab}}^{\text{AS}}$:

$$\frac{U_{l_2^L}}{w_2^L} = (1 + \chi_{\text{lab}}^{\text{AS}})\left(\frac{U_{l_2^H}}{w_2^H}\right).$$

3. *Intertemporal consumption wedge*: a "tax" on relocating consumption from the first to the second period, i.e., a savings tax. For each unit of reduction in the first-period consumption, the second-period consumption only increases by $1 - \chi_{\text{cons}}^{\text{LC}}$:

$$U_{c_1} = (1 - \chi_{\text{cons.}}^{\text{LC}})\mathbf{E}\{U_{c_2}\}.$$

4. *Consumption dispersion wedge*: a "tax" on relocating consumption from the high- to low-wage state in the second period. For each unit of consumption reduction in the high-wage state, the low-wage state consumption only increases by $1 - \chi_{\text{cons.}}^{\text{AS}}$:

$$U_{c_2^H} = (1 - \chi_{\text{cons}}^{\text{AS}})U_{c_2^L}.$$

5. *Consumption-labor wedge*: a labor income "tax" in the first period on converting labor

to consumption. For each unit of earnings, consumption only rises by $1-\chi_{\text{cons,lab}}$:

$$-\frac{U_{l_1}}{w_1} = (1-\chi_{\text{cons.,lab}})U_{c_1}.$$

The intertemporal and labor (consumption) dispersion wedges capture inefficiencies in the allocation of labor (consumption) over time and across idiosyncratic states, respectively. The consumption-labor wedge measures distortions in the levels of consumption and labor. If all five wedges are zero and the resource constraint (3) is satisfied, the allocation is efficient.

Table 10 summarizes the formal definitions of wedges and presents their signs under various equilibria in the columns denoted "Wedge Signs". In the laissez-faire equilibrium, there are no distortions over time or between consumption and labor, and hence the intertemporal and consumption-labor wedges are zero. However, due to missing insurance markets, the distributions of labor and consumption are inefficient across idiosyncratic states. Specifically, the labor and consumption dispersion wedges are both positive, implying too little labor and too much consumption in the high-wage state relative to the low-wage state.

Table 10: Wedges in the Two-Period Model

		I. Wedge Signs				II. When Progressivity ↑		
Wedges	Definition	Efficient	LF	AIT	LIT	AIT	LIT	Compare
A. Labor Distribution								
Intertemporal	$\frac{U_{l_1}/w_1}{\mathbf{E}\{U_{l_2}/w_2\}}-1$	0	0	$\gtreqless 0$	0	↑	→	LIT>AIT
Dispersion	$\frac{U_{l_2^L}/w_2^L}{U_{l_2^H}/w_2^H}-1$	0	>0	>0	>0	↓	⇊	LIT>AIT
B. Consumption Distribution								
Intertemporal	$1-\frac{U_{c_1}}{\mathbf{E}\{U_{c_2}\}}$	0	0	0	0	→	→	LIT=AIT
Dispersion	$1-\frac{U_{c_2^H}}{U_{c_2^L}}$	0	>0	>0	>0	⇊	↓	LIT<AIT
C. Levels of Consumption and Labor								
Consumption-Labor	$1+\frac{U_{l_1}/w_1}{U_{c_1}}$	0	0	$\gtreqless 0$	$\gtreqless 0$	↑	⇈	LIT<AIT

Notes: "LF" denotes the laissez-faire equilibrium. For Column Group II, results are based on the parameterized two-period model in Appendix B.1 with $\mathbf{E}\{w_2\}/w_1 = 1.109$ and $\Delta = 0.303$. The symbols ↑, ↓, and → indicate an increase, decrease, or no change, respectively, while double arrows represent stronger effects.

A progressive AIT distorts labor supply decisions over time as the tax rate is higher in periods with more earnings, and thus the intertemporal labor wedge may be positive or negative. Proposition 3 states that a LIT introduces no such distortion, and therefore, the intertemporal labor wedge is zero. Both progressive annual and lifetime income taxes may provide partial insurance against wage risks. However, because insurance remains incomplete and labor supply is distorted, the labor and consumption dispersion wedges are positive. The consumption-labor wedge may be positive or negative, depending on the overall tax level.

B.1 Parameterization

To illustrate the heterogeneous effects of lifetime and annual income taxes on the wedges and identify their key determinants and consequences for welfare, we parameterize the two-period model and solve it numerically under various fiscal constitutions. The period utility function is additively separable between consumption and labor and given by (1). We choose $\sigma = 1.5$ and $\eta = 0.5$ as in our quantitative analysis, and the parameter ψ is normalized to 1 for simplicity. The first-period wage w_1 is deterministic, but there is idiosyncratic wage risk in the second period. The expected wage is denoted by $\mathbf{E}\{w_2\} = 1$, but it can take the values $w_2^H = 1 + \Delta$ or $w_2^L = 1 - \Delta$ with equal probability. Therefore, $\mathbf{E}\{w_2\}/w_1 = 1/w_1$ is the expected wage growth from the first to the second period, and Δ measures the extent of idiosyncratic wage risk (its standard deviation) in the second period.

Following Bénabou (2002) and Heathcote et al. (2017), and consistent with our quantitative analysis, we adopt a two-parameter functional form for both the AIT and LIT:

$$T_X(y) = y - (1 - \tau_X)y^{1-\mu_X}, \ X \in \{A, L\},$$

where y denotes pre-tax earnings, and μ_X and τ_X are policy parameters governing the progressivity and the level of income tax. The subscript A and L represent annual and lifetime income taxes. For the status quo AIT, the tax progressivity and level are set to 0.137 and 0.105, as in our quantitative analysis. Note that Proposition 4 and Corollary 5 in Appendix A.4 demonstrate that with the aforementioned assumptions in the two-period model, all wedges—except for the consumption-labor wedge—remain invariant with respect to the tax *level* parameter τ, under both the AIT and LIT systems.

B.2 LIT vs. AIT: Fixed Progressivity

We now examine the effects of replacing the status quo AIT with a LIT of the same progressivity, highlighting the key role of the wage process in shaping the results. The tax level parameter for the LIT is determined endogenously to match the tax revenue of the status quo policy. The top four panels of Figure 8 compare distortions introduced by both tax systems, measured by the absolute values of wedges defined in Table 10. The horizontal axis $\mathbf{E}\{w_2\}/w_1$ represents expected wage growth across the two periods, while the vertical axes Δ indicate the level of wage risk in the second period. A positive value (blue color) indicates a smaller wedge under LIT than under AIT, implying less distortions by LIT.

The intertemporal labor wedge is zero under LIT (as in the efficient allocation), whereas under AIT, the wedge may be positive or negative, indicating distortions (top-left panel). LIT also leads to a smaller labor dispersion wedge (i.e., a more efficient labor distribution within the second period) compared to AIT, except when wage growth is high and wage risk

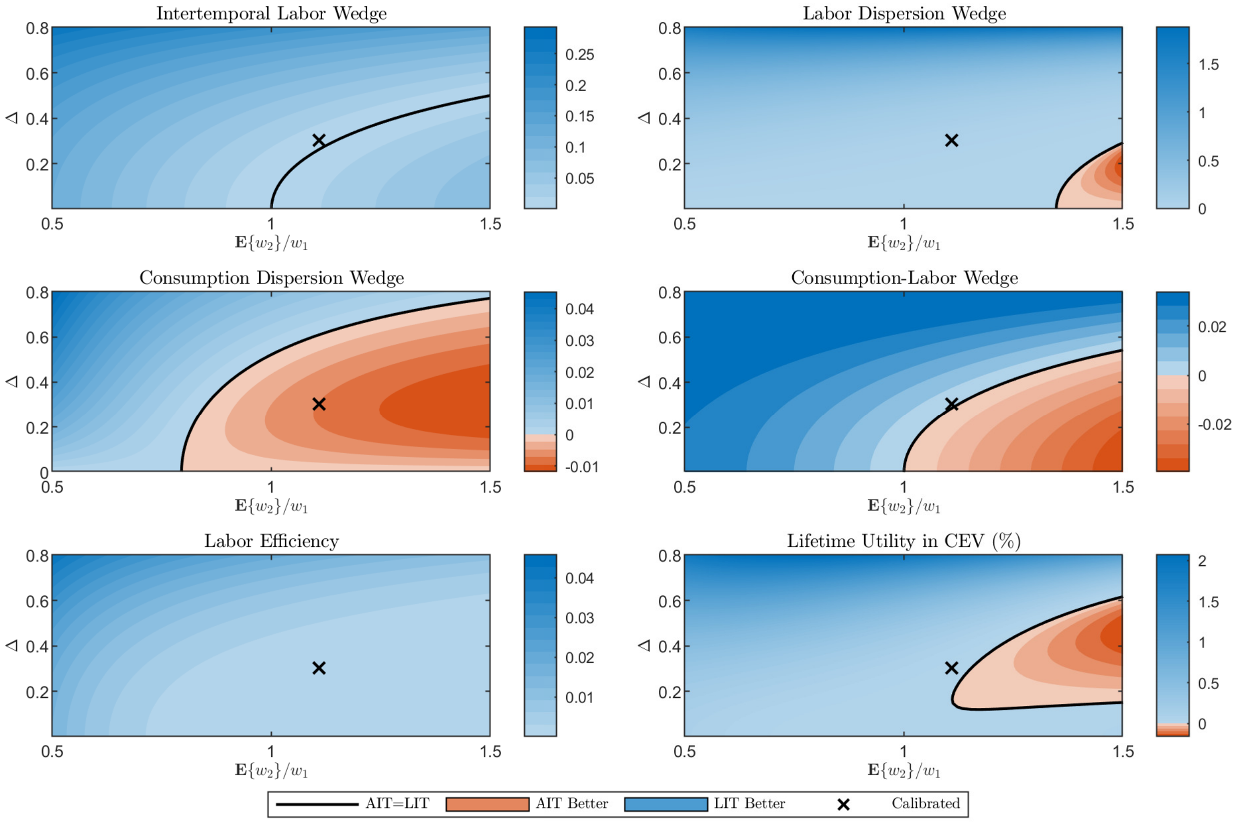

Figure 8: LIT vs. AIT: Fixed Progressivity

Notes: This figure compares LIT and AIT under the same tax progressivity. The horizontal axis measures average wage growth between two periods, while the vertical axis measures wage risk in the second period. The top four panels display differences in the absolute values of wedges; the bottom-left and bottom-right panels show differences in labor efficiency and lifetime utility. A positive value (blue areas) indicates that LIT has smaller wedges than AIT, higher labor efficiency, or greater lifetime utility—by that magnitude. A negative value (red areas) implies the opposite. Tax progressivity is set at $\mu_A = \mu_L = 0.137$. The tax level parameter is $\tau_A = 0.105$ and for the LIT, τ_L is endogenously determined to generate the same tax revenue as AIT. The "**x**" symbol in the plots represent the calibrated wage growth and wage risk, based on the differences in trend wage and the variance of log wages between age 35 and 55 in the PSID, and thus signifies the parameterization of $\mathbf{E}\{w_2\}/w_1$ and Δ we view as most informative.

is small. Even when AIT outperforms LIT (red area), the advantage is small. In contrast, when LIT performs better (blue area), the difference is more substantial.

The middle-left panel illustrates the consumption dispersion wedge, which measures the degree of consumption insurance between the low- and high-wage states. AIT provides better consumption insurance when wage growth is significant and wage risk is moderate. The intuition is that when average earnings increase between the two periods, AIT effectively mimics an age-dependent tax with a higher tax rate in the second period, thereby reducing second-period labor supply. Since wage risk is concentrated in the second period, this mechanism helps mitigate its impact on household consumption. The middle-right panel depicts the consumption-labor wedge, which measures the efficiency trade-off between the overall level of consumption and labor supply. AIT dampens labor supply less (i.e., introduces a smaller wedge) when wage growth is substantial and wage risk is small.

As indicated by the top two panels, LIT generally introduces less distortions to labor supply decisions. Consequently, aggregate labor efficiency, defined as the ratio of total earnings to total hours worked, is consistently higher under LIT than AIT, as shown in the bottom-left panel of Figure 8. In contrast, AIT tends to offer better consumption insurance and may have a smaller depressive effect on overall labor supply.

Given the respective advantages and disadvantages of LIT and AIT, which one induces a higher level of social welfare is a quantitative question. The bottom-right panel of Figure 8 illustrates the difference in lifetime utility under LIT and AIT in terms of consumption-equivalent variation. LIT outperforms AIT when wage growth is modest or negative, regardless of wage risk, as well as when wage growth is significant and wage risk is either small or large (blue area). However, AIT can surpass LIT in cases where wage growth is significant and wage risk is moderate (red area).

The "**x**" symbol in Figure 8 marks the wage process with $\mathbf{E}\{w_2\}/w_1 = 1.109$ and $\Delta = 0.303$, calibrated to match the differences in trend wage and log-wage variance between ages 35 and 55 in the PSID. With this calibration, the labor efficiency gain from LIT outweighs the consumption insurance loss, resulting in higher lifetime utility under LIT than AIT. However, this experiment corresponds to a shift from a semi-lifetime income tax to a LIT. In reality, the status quo policy operates at an annual frequency, various types of idiosyncratic wage shocks are resolved gradually over the life cycle, and the life-cycle wage profile is hump-shaped, not monotonic. Thus, assessing the consequences of a LIT requires a full life-cycle model that captures these features, motivating the full quantitative analysis in the main text.

B.3 LIT vs. AIT: Varying Progressivity

Column group II of Table 10 summarizes the effects of increasing tax progressivity on the wedges, both under LIT and AIT. Figure 9 illustrates how the wedges (in absolute value) vary with tax progressivity quantitatively. For this experiment, the wage process is the same as that marked by "**x**" in Figure 8, and the tax level parameters are determined endogenously to match the tax revenue of the status quo policy.

Increasing tax progressivity enhances labor efficiency under LIT, as the intertemporal wedge remains zero (solid blue line, top-left panel) and the labor dispersion wedge moves closer to zero (solid blue line, top-right panel). This reduction in the labor dispersion wedge occurs because a more progressive income tax improves consumption equality, thereby narrowing the gap in marginal utility of consumption between high- and low-wage states. As

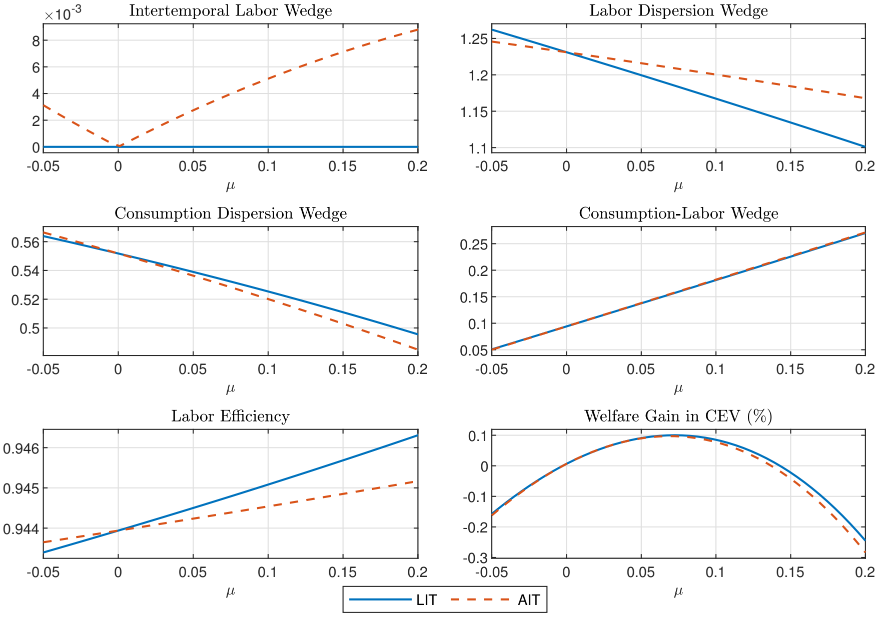

Figure 9: LIT vs. AIT: Varying Progressivity

Notes: This figure shows how wedges, labor efficiency, and lifetime utility change with tax progressivity μ under AIT and LIT. The top four panels display the absolute values of wedges, the bottom-left panel shows labor efficiency, and the bottom-right panel presents the change in lifetime utility relative to the status quo AIT. The tax level parameters τ_A and τ_L vary with μ to match the tax revenue under the status quo policy.

a result, labor supply increases in the high-wage state and decreases in the low-wage state when financial markets are incomplete. For AIT, greater progressivity also reduces the labor dispersion wedge (dashed red line, top-right panel) in this example, but less effectively than LIT.[54] Moreover, a more progressive AIT can introduce greater intertemporal labor distortions (dashed red line, top-left panel, when $\mu > 0$), thereby reducing labor efficiency.

The middle-left panel shows that the primary benefit of a more progressive income tax is the reduction in the consumption dispersion wedge, i.e., better consumption insurance, leading to a more efficient distribution of consumption. The main drawback is a larger consumption-labor wedge, which further depresses labor supply relative to consumption, as shown in the middle-right panel. AIT appears to manage this trade-off more effectively by delivering a greater improvement in consumption insurance with a relatively smaller increase in the consumption-labor wedge. However, a more progressive LIT has the additional advantage of enhancing labor efficiency more, as demonstrated in the bottom-left panel.

---

[54] It is worth noting that this reduction under AIT is not guaranteed. For instance, if Δ is doubled while other factors remain constant, the labor dispersion wedge increases with AIT progressivity. In contrast, under LIT, this wedge continues to decline with greater progressivity. In general, labor efficiency moves either more positively or less adversely with tax progressivity under a LIT than an AIT.

C Supplementary Results for Quantitative Analysis

In this section, we provide supplementary results to the quantitative analysis in the main text.

C.1 Additional Calibration Results

C.1.1 Survival Probability and Population Structure

Figure 10 plots the life-cycle profile of conditional survival probabilities (left panel) and the implied population shares by age (right panel).

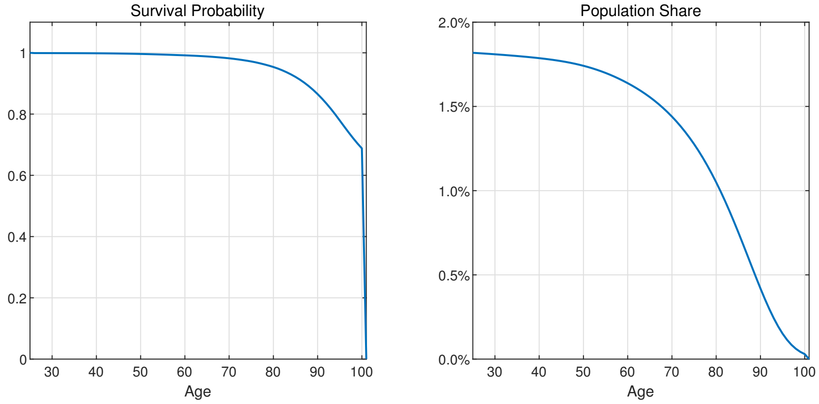

Figure 10: Survival Probability and Population Structure

Notes: This figure plots the life-cycle profiles of survival probabilities (left panel), based on the Social Security Administration Period Life Table 2013, and the corresponding population shares (right panel).

C.1.2 Wage Trend

Figure 11 displays the deterministic log-wage trends over the life cycle, estimated from the 1999-2017 PSID data.

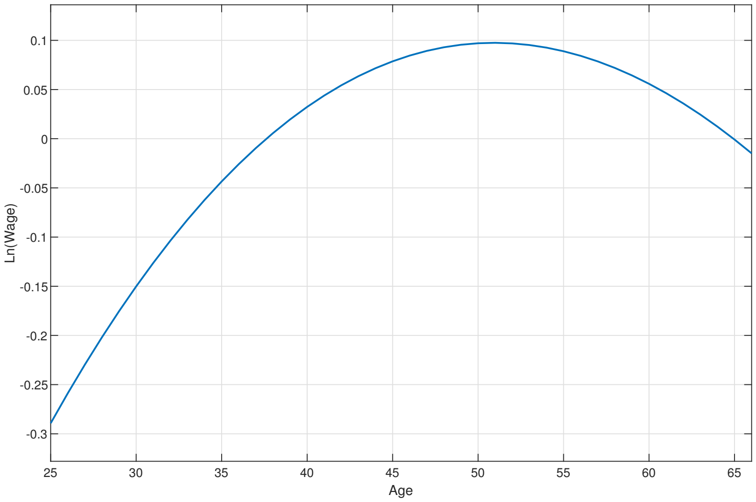

Figure 11: Life-Cycle Wage Trend

Notes: This figure plots the deterministic log-wage trend over household life cycle estimated from the PSID data.

C.1.3 Idiosyncratic Wage Process

Figure 12 compares the patterns of wage persistence observed in the data with those implied by the estimated wage process.[55] The left panel shows that wage persistence depends on both the previous wage and the shock received: persistence is higher when the previous wage and the shock are of similar ranks. In other words, persistence is greater when high (low) wage earners experience more positive (negative) shocks. The middle panel shows that the estimated wage process captures this pattern in the data well. Finally, the right panel reveals that wage persistence increases modestly with age, a pattern closely mirrored by the estimated wage process.

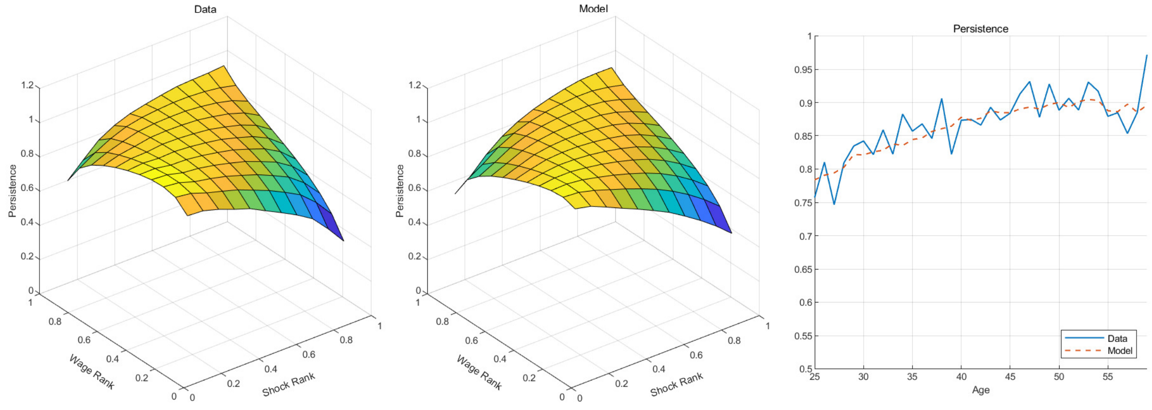

Figure 12: Model vs. Data: Wage Persistence

Notes: This figure plots the wage persistence implied by the PSID 1968-1993 data and the estimated wage process. The left and middle panels present how the wage persistence varies with the rank of previous wage and the rank of wage shock received, in the data and in the model, respectively, while the right panel plots the life-cycle profiles of wage persistence.

Figure 13 plots the distribution of one-year residual log-wage growth in both the data and the model, compared to a normal distribution with the same variance as the data. The model closely matches the shape of the data distribution, with both exhibiting higher kurtosis than a normal distribution. Figures 14 and 15 further demonstrate that the model also captures the dispersion, skewness, and kurtosis of the distribution of log-wage growth conditional on previous wage and age.[56] In both the data and the estimated wage process, using both standard and robust measures, the conditional dispersion of wage growth is larger at the bottom and top of the wage distribution (left panels of Figure 14), and the conditional

---

[55] Since we do not observe the persistent and transitory components separately in the data, we can only compare wage persistence. This is obtained by estimating the quantile function of residual log-wage, conditional on the previous residual log-wage and age, $\hat{w}_j = Q_w(\hat{w}_{j-1}, j, u_{w,j})$, and then computing $\partial Q_w(\hat{w}_{j-1}, j, u_{w,j})/\partial \hat{w}$, following the same technique used to estimate Q_z in Arellano et al. (2017). The wage persistence implied by the model is derived by applying the same procedures to simulated data from the estimated wage process.

[56] We report the standard deviation, skewness, and kurtosis, as well as their corresponding robust measures

skewness (kurtosis) decreases (increases) with previous wage (middle and right panels of Figure 14). Conditional dispersion decreases modestly with age (left panels of Figure 15), while conditional skewness and kurtosis remain roughly stable over the life cycle (middle and right panels of Figure 15).

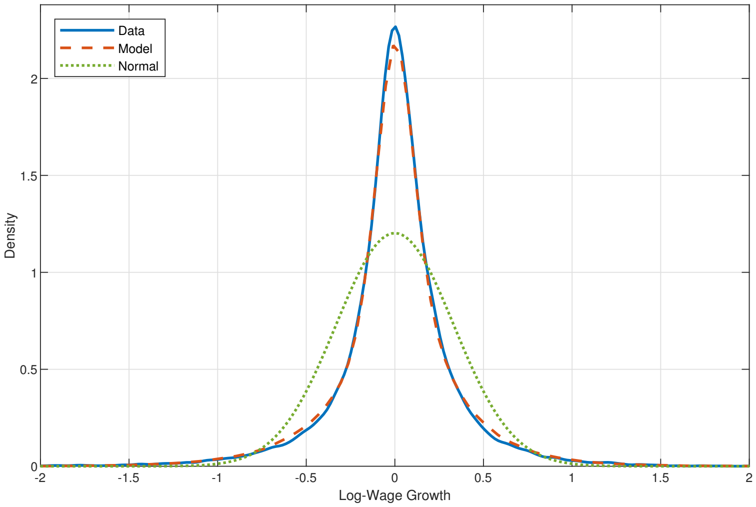

Figure 13: Model vs. Data: Distribution of Log-Wage Growth

Notes: This figure plots the distribution of one-year residual log-wage growth in the PSID 1968-1993 data (blue solid line), and that implied by the estimated wage process (red dashed line), alongside a normal distribution of the same data variance (green dotted line).

When discretizing the estimated wage process following the method in De Nardi et al. (2019), we use 13 grid points per age for the persistent component, and 6 grid points for the transitory component, balancing the accuracy of the approximation and computation cost. To capture the higher wage inequality in the 2000s, we scale up the grids for the wage process estimated from the 1968–1993 PSID annual data to match the life-cycle profile of the variance of log-wage in the 1999-2017 PSID biennial data.

Figures 16, 17, 18, and 19 are counterparts to Figures 12, 13, 14, and 15, but for the discretized wage process. These figures show that the discretized wage process captures the features of the data reasonably well, with the scaling up of the grids having almost negligible effects, except for the dispersion of wage growth.

As a test of how well our model matches the 1999–2017 PSID biennial data, we also

---

based on quantiles, as in Arellano et al. (2017). In particular, we define the robust measures as follows:

$$\begin{aligned} \text{Std(R)} &\equiv Q(1-\alpha_1) - Q(\alpha_1), \\ \text{Skewness(R)} &\equiv \frac{Q(1-\alpha_1) + Q(\alpha_1) - 2Q(1/2)}{Q(1-\alpha_1) - Q(\alpha_1)}, \quad \text{(Kelley's skewness)}, \\ \text{Kurtosis(R)} &\equiv \frac{Q(1-\alpha_1) - Q(\alpha_1)}{Q(1-\alpha_2) - Q(\alpha_2)}, \quad \text{(Crow–Siddiqui kurtosis)}, \end{aligned}$$

where Q is the quantile function, $\alpha_1 = 1/12$, and $\alpha_2 = 2/12$.

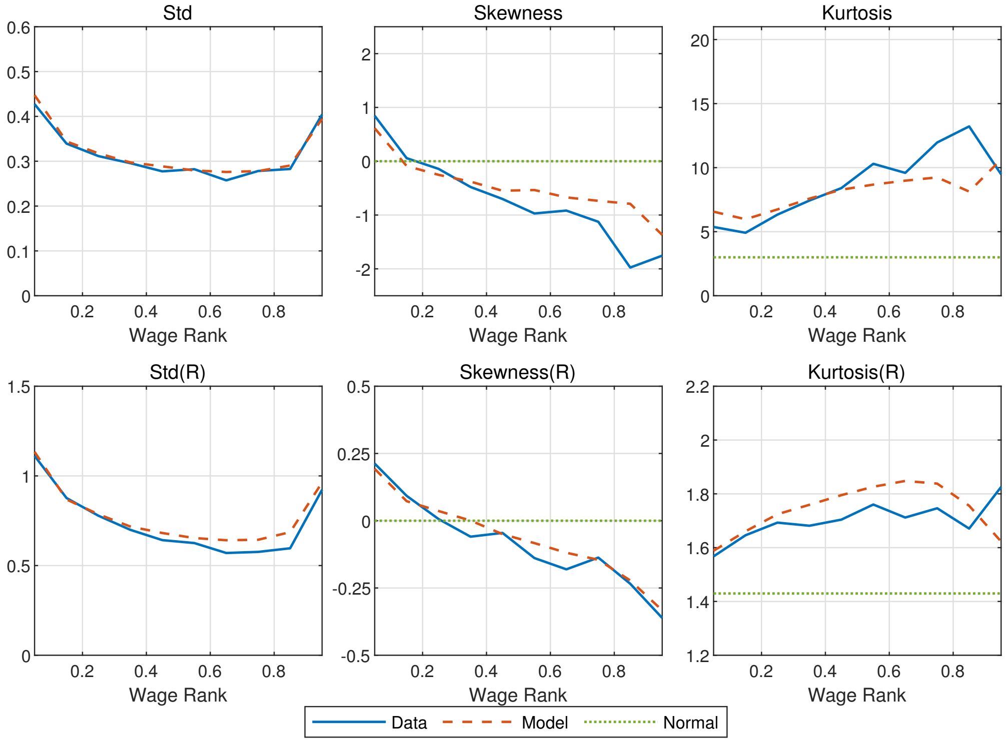

Figure 14: Model vs. Data: Distribution of Log-Wage Growth by Previous Wage Rank

Notes: This figure plots the dispersion, skewness, and kurtosis of the distribution of one-year residual log-wage growth conditional on the rank of previous wage in the PSID 1968-1993 data (blue solid lines) and implied by the estimated wage process (red dashed lines), alongside those implied by a normal distribution (green dotted lines). The top panels use the standard definitions of standard deviation, skewness, and kurtosis, whereas the bottom panels use corresponding robust measures based on quanitles as in Arellano et al. (2017).

compare the distributions of *two-year* residual wage growth. Figure 20 shows that our model matches this distribution in the data reasonably well, despite not targeting it in the estimation.[57]

---

[57] We can compute only the distribution of two-year wage growth for this time period due to the biennial nature of the PSID. The plot excludes measurement errors from the model results as the shape of the measurement error distribution is unknown.

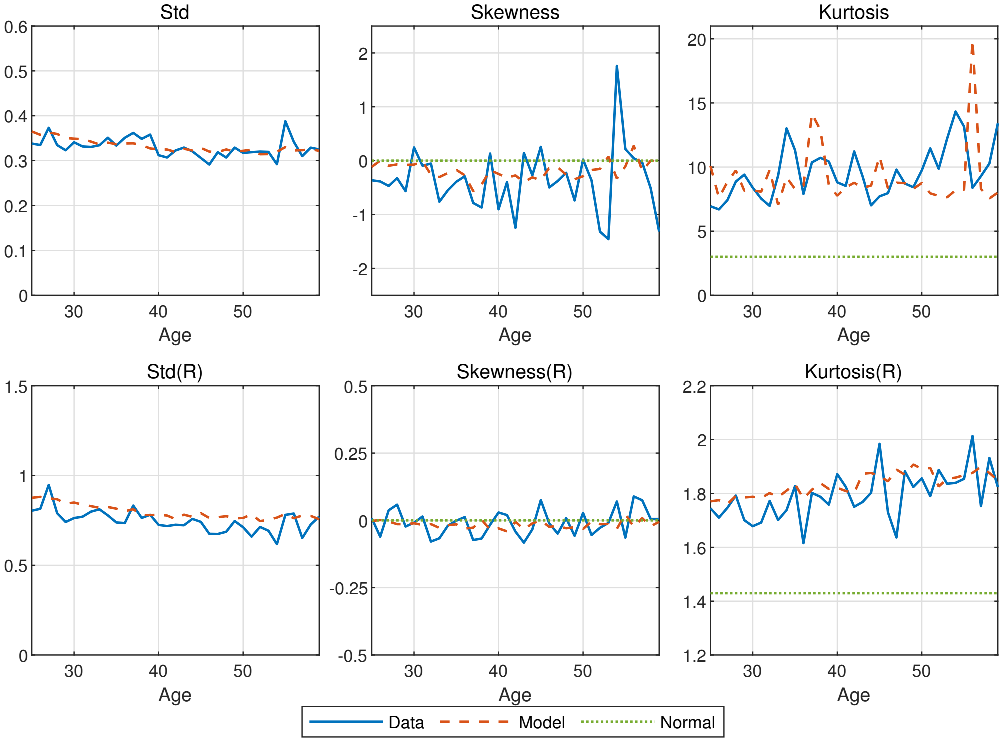

Figure 15: Model vs. Data: Distribution of Log-Wage Growth by Age

Notes: This figure plots the dispersion, skewness, and kurtosis of the distribution of one-year residual log-wage growth conditional on age in the PSID 1968-1993 data (blue solid lines) and implied by the estimated wage process (red dashed lines), alongside those implied by a normal distribution (green dotted lines). The top panels use the standard definitions of standard deviation, skewness, and kurtosis, whereas the bottom panels use corresponding robust measures based on quanitles as in Arellano et al. (2017).

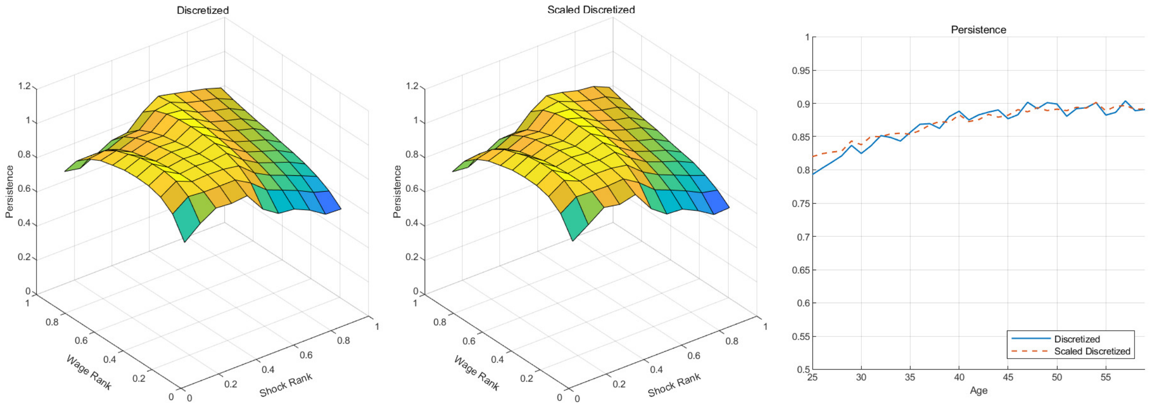

Figure 16: Discretized Wage Process: Wage Persistence

Notes: This figure plots the wage persistence implied by the discretized wage process based on PSID 1968–1993 and its scaled version, which matches the wage variances in PSID 1999–2017. The left and middle panels show how wage persistence varies with the rank of the previous wage and the rank of the wage shock received, as implied by each process, respectively. The right panel plots the life-cycle profiles of wage persistence.

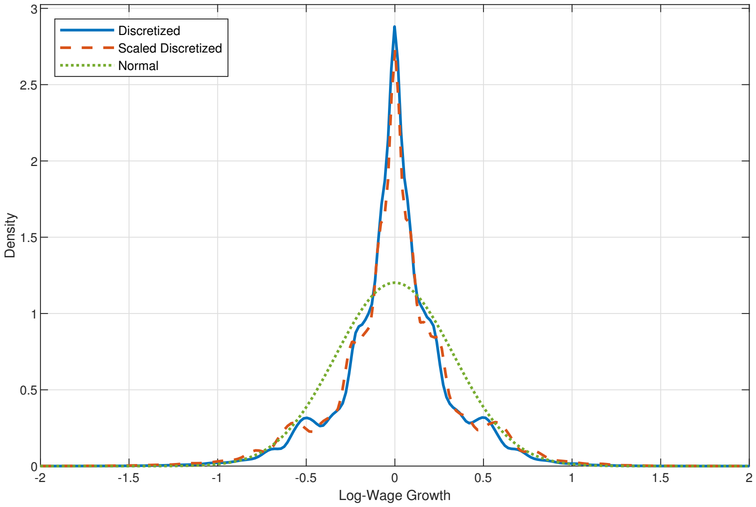

Figure 17: Discretized Wage Process: Distribution of Log-Wage Growth

Notes: This figure plots the distribution of one-year residual log-wage growth implied by the discretized wage process based on PSID 1968–1993 (blue solid line) and its scaled version (red dashed line), which matches the wage variances in PSID 1999–2017, alongside a normal distribution of the same data variance in PSID 1968-1993 (green dotted line).

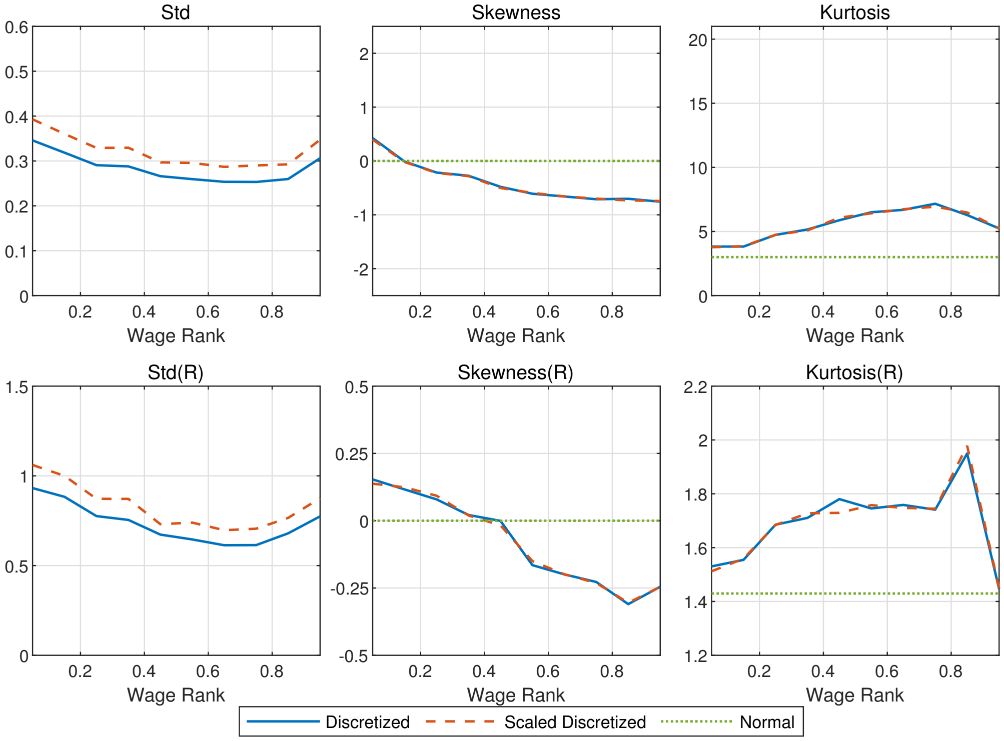

Figure 18: Discretized Wage Process: Distribution of Log-Wage Growth by Previous Wage Rank

Notes: This figure plots the dispersion, skewness, and kurtosis of the distribution of one-year residual log-wage growth conditional on the rank of previous wage implied by the discretized wage process based on PSID 1968–1993 (blue solid lines) and its scaled version (red dashed lines), which matches the wage variances in PSID 1999–2017, alongside those implied by a normal distribution (green dotted lines). The top panels use the standard definitions of standard deviation, skewness, and kurtosis, whereas the bottom panels use corresponding robust measures based on quanitles as in Arellano et al. (2017).

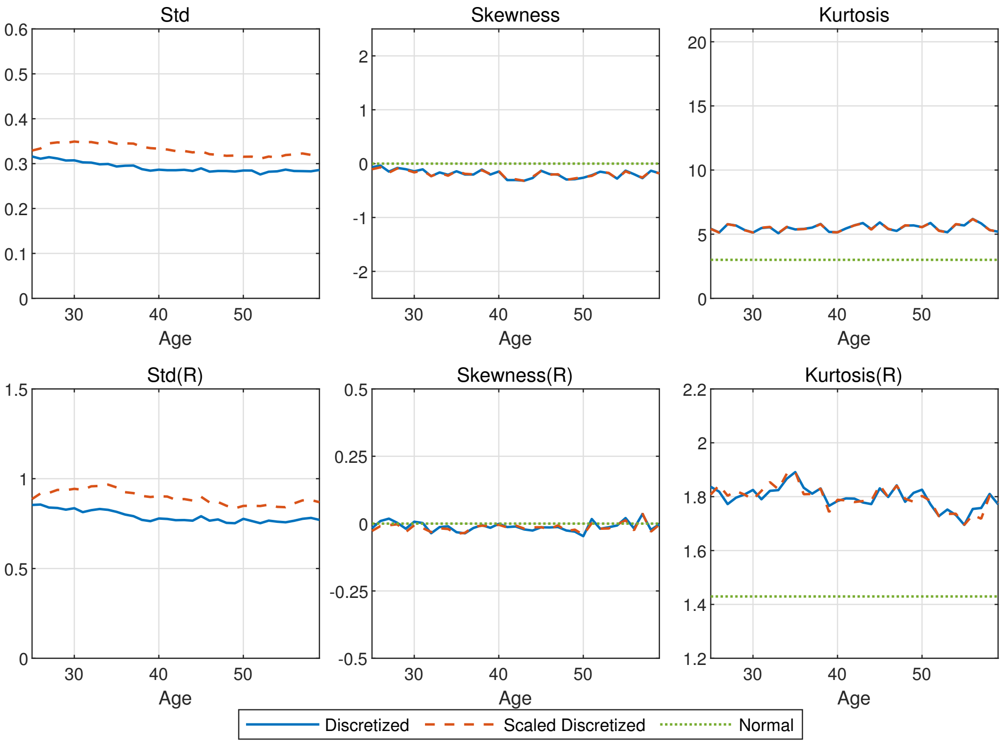

Figure 19: Discretized Wage Process: Distribution of Log-Wage Growth by Age

Notes: This figure plots the dispersion, skewness, and kurtosis of the distribution of one-year residual log-wage growth conditional on age implied by the discretized wage process based on PSID 1968–1993 (blue solid lines) and its scaled version (red dashed lines), which matches the wage variances in PSID 1999–2017, alongside those implied by a normal distribution (green dotted lines). The top panels use the standard definitions of standard deviation, skewness, and kurtosis, whereas the bottom panels use corresponding robust measures based on quanitles as in Arellano et al. (2017).

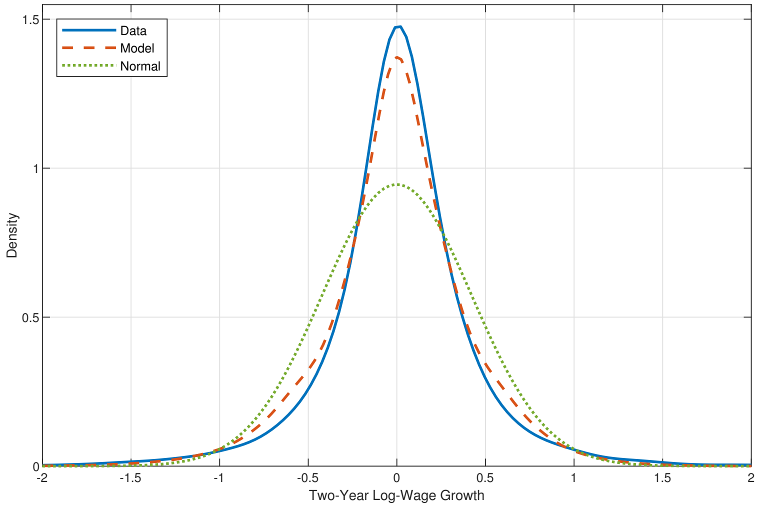

Figure 20: Model vs. Data: Distribution of Two-Year Log-Wage Growth

Notes: This figure plots the distribution of two-year residual log-wage growth in the PSID 1999-2017 biennial data (blue solid line) and in the baseline life-cycle model (red dashed line), alongside a normal distribution of the same data variance (green dotted line). Model results exclude measurement errors.

C.1.4 Initial Net Worth Distribution

Figure 21 compares the empirical cumulative distribution function of initial net worth in the data with that implied by the discretized version.

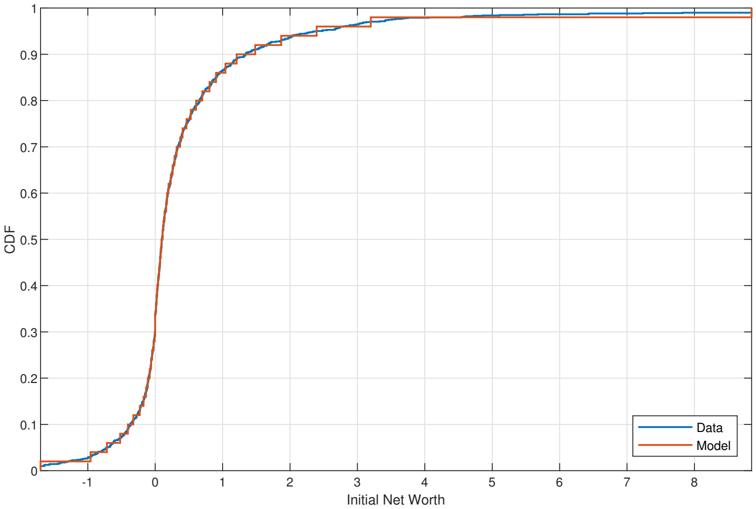

Figure 21: Cumulative Distribution Function of Initial Net Worth

Notes: This figure plots the cumulative distribution function of net worth among age-25 households in the PSID 1999–2017 data (blue line), alongside that implied by the approximated initial net worth grid in the model (red line).

C.1.5 Share of Borrowing Constrained

Figure 22 presents the share of borrowing-constrained households over the life cycle in the status quo economy. Only a negligible share of households are borrowing-constrained before retirement, whereas the share rises sharply around age 90 as long-lived households deplete their savings and rely solely on retirement benefits.

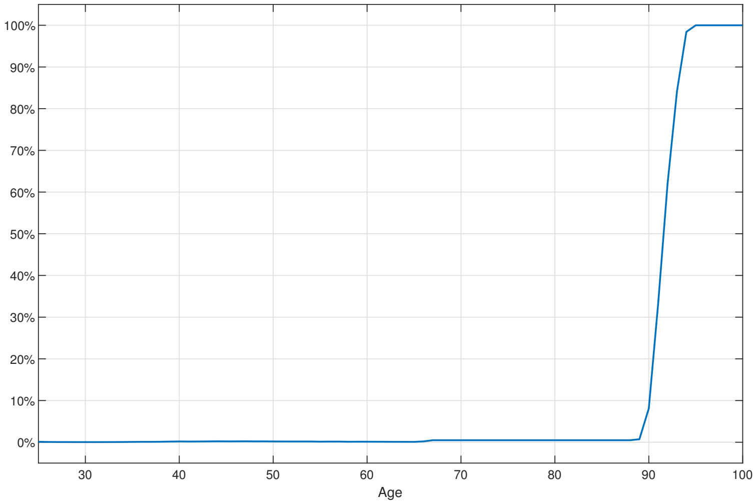

Figure 22: Share of Borrowing Constrained Households

Notes: This figure plots the percentage of borrowing constrained households in the status quo economy.

C.1.6 Wealth Inequality and Consumption Insurance

Table 11 reports statistics on the net worth distribution in the PSID data and our benchmark model. For comparison, we also include results from De Nardi et al. (2019).

Table 11: Model vs. Data: Wealth Distribution

	Wealth	Wealth Shares of Top Groups					
	Gini	1%	5%	20%	40%	60%	80%
Data	0.82	0.24	0.49	0.81	0.96	1.02	1.03
Model	0.61	0.07	0.23	0.59	0.84	0.97	1.02
Model DFP	0.61	0.07	0.25	0.61	0.85	0.96	0.999

Notes: This table reports statistics of the net worth distribution for households aged 25-60 in the PSID 1999-2017 and in the status quo model economy. "Model DFP" corresponds to model results in De Nardi et al. (2019) with nonlinear earnings process.

Table 12 reports measures of consumption insurance implied by our model alongside those estimated in the previous empirical literature. We measure consumption insurance in our model using the insurance coefficient defined by Blundell et al. (2008) (BPP) and Kaplan and Violante (2010), but applied to the wage shocks modeled in our paper, rather than earnings shocks as in BPP:

$$1 - \frac{\text{cov}(\Delta c_j, x_j)}{\text{var}(x_j)}.$$

Here Δc_j denotes the log consumption growth from age $j-1$ to age j that is not explained by observables (e.g., age), and x_j represents the shock at age j, which can be either the innovation to the persistent or transitory log-wage component. Because the shocks are observable in the model, we can compute these insurance coefficients directly. The results are reported in the "Model True" column of Table 12. The insurance coefficient is 0.67 for persistent shocks, meaning that 33% of persistent wage shocks pass through to consumption. In comparison, the insurance coefficient is 0.98 for transitory shocks, implying that only 2% of transitory wage shocks pass through to consumption.

Table 12: Consumption Insurance Coefficients

	Model True	Model BPP		Data BPS	Data ABB	Data BPP
	Shocks	Wage	Earnings	Wage	Earnings	Earnings
Persistent	0.67	0.65	0.71	0.68	0.6–0.7	0.69
Transitory	0.98	0.92	0.94	1.14	N/A	0.94

Notes: This table reports model true consumption insurance coefficients computed from wage shocks, model BPP coefficients computed from simulated wage and earnings data, and empirical coefficients from the literature. "BPP", "BPS", and "ABB" correspond to Blundell et al. (2008), Blundell et al. (2016), and Arellano et al. (2017), respectively.

Unfortunately, it is difficult to find exact empirical counterparts for these coefficients in the literature. Existing empirical studies either assume a canonical permanent-transitory linear process (Blundell, Pistaferri, and Saporta-Eksten 2016), estimate consumption insurance against earnings shocks rather than wage shocks (Arellano et al. 2017), or both (Blundell et al. 2008). However, based on the available empirical evidence, the overall conclusion from Table 12 emerges that the degree of consumption insurance implied by our model is broadly consistent with this existing empirical evidence.

We now describe how the estimates for the consumption insurance coefficients in the last five columns of the table are derived. The "Wage" subcolumn of the "Model BPP" column reports insurance coefficients estimated from model-simulated wage data using the formulas in Kaplan and Violante (2010):

$$\text{Persistent: } 1 - \frac{\mathrm{cov}(\Delta c_j, \widehat{w}_{j+1} - \widehat{w}_{j-2})}{\mathrm{cov}(\Delta \widehat{w}_j, \widehat{w}_{j+1} - \widehat{w}_{j-2})}, \quad \text{Transitory: } 1 - \frac{\mathrm{cov}(\Delta c_j, \Delta \widehat{w}_{j+1})}{\mathrm{cov}(\Delta \widehat{w}_j, \Delta \widehat{w}_{j+1})}, \tag{27}$$

where $\widehat{w}_j$ denotes the residual log wage at age j. These formulas identify the true insurance coefficients *only under the canonical permanent–transitory log-linear wage process*, and thus does not apply exactly in our model. Nevertheless, the estimates based on model-simulated data are quite close to the true values, with a somewhat larger bias for transitory shocks. The empirical estimates from Blundell et al. (2016) (fourth column of the table, "Data BPS") can be seen as providing the best approximate empirical counterpart to our "Model BPP/Wage" results and are therefore informative about the true consumption insurance coefficients. Using PSID data from 1999 to 2009, they estimate a pass-through rate of 32% to consumption of permanent wage shocks of men, implying an associated consumption insurance coefficient of 0.68, which is very close to that implied by our model. They also find a slightly negative pass-through rate of −14% from transitory shocks, whereas our model implies a near-zero pass-through rate.

Another related empirical literature to which we can compare our model insurance coefficients against treats household earnings as exogenous and estimates consumption insurance against these earnings shocks. Since in our model earnings are endogenous, the results are not directly comparable to those in this literature. Nevertheless, we can view the "consumption insurance coefficients against earnings shocks" estimated from our model-simulated data as informative moments and compare them to the empirical estimates. The "Earnings" sub-column of the "Model BPP" column (colum 3 of the table) reports these coefficients, computed using the same formulas in (27) but based on residual log earnings rather than residual log wages.[58] The estimated model-implied insurance coefficient is 0.71 for persistent

---

[58] When applied to consumption insurance against earnings shocks, these formulas rely on the assumption of a canonical permanent-transitory linear earnings process, which also does not apply in our model, and therefore there is no "Model True" consumption insurance coefficient for shocks to earnings.

"earnings shocks" and 0.94 for transitory "earnings shocks".

Arellano et al. (2017) assume a nonlinear, non-Gaussian, history- and age-dependent earnings process and estimate flexible consumption responses to earnings shocks using PSID 1999-2009 data. Their estimates of consumption insurance against persistent earnings shocks range from 0.6 to 0.7, depending on household states, close to the estimate for our model of 0.71. Arellano et al. (2017) do not report estimates for transitory earnings shocks. Blundell et al. (2008), assuming a canonical permanent-transitory linear earnings process, find a pass-through rate of 31% from permanent pre-tax earnings shocks to consumption, corresponding to a consumption insurance coefficient of 0.69, which is also close to our model counterpart. Their estimate of the coefficient against transitory earnings shocks is 0.94, identical to the one in our model.

There are three main reasons why our model is able to match well the high degree of consumption insurance observed in the data, especially against persistent shocks. First, we focus on wage shocks rather than earnings shocks, and endogenous labor supply provides an additional mechanism for insurance against idiosyncratic risk. Second, as documented by De Nardi et al. (2019) and Guvenen et al. (2024), the rich wage dynamics (earnings dynamics in their papers) we model, including the lower persistence of shocks among young households, tend to increase partial consumption insurance. As Kaplan and Violante (2010) show, modeling a persistent instead of a fully permanent component substantially increases consumption insurance in the model. Finally, we allow for more generous borrowing limits early in life, motivated by the fact that a notable fraction of young households have significantly negative net worth, than many previous model-based analyses that assume a tight borrowing limit at zero.

C.2 Welfare Decomposition

In this section, we explain how welfare changes in consumption-equivalent variations (CEV) and the decomposition of welfare changes into level, age, and distribution effects of consumption and labor are calculated in the main text. The notations are independent from the main text or other sections of the appendix.

Let $(\mathbf{c}, \mathbf{h})$ denote the state-contingent plan for household consumption and labor, and let $W(\mathbf{c}, \mathbf{h})$ represent the social welfare under this state-contingent plan. Consider a change in consumption and labor allocations from $(\mathbf{c}^0, \mathbf{h}^0)$ to $(\mathbf{c}^1, \mathbf{h}^1)$. The welfare effect of this allocation change, expressed in consumption-equivalent variation, CEV, is defined by the following equation:

$$W((1 + CEV)\mathbf{c}^0, \mathbf{h}^0) = W(\mathbf{c}^1, \mathbf{h}^1).$$

That is, CEV is the percentage change in lifetime consumption (i.e., consumption at all ages and all contingent states) required to generate a change in social welfare equivalent to that induced by the allocation change. If CEV is positive (negative), the allocation change is welfare-improving (welfare-reducing).

When the allocation change reflects the total effects of a tax reform, the corresponding CEV measures the total welfare gain from the reform. In the welfare decomposition exercise, we divide the allocation change induced by the reform into three types of changes, for both consumption and labor: a change in the average level, a change in the age-profile, and the remaining distributional changes. Using consumption as an example, these three types of changes are defined as follows:

1. Consumption level: $\mathbf{c}^1 = \frac{\bar{\mathbf{c}}^1}{\bar{\mathbf{c}}^0}\mathbf{c}^0$, $\mathbf{h}^1 = \mathbf{h}^0$, where $\bar{\mathbf{c}}^0$ and $\bar{\mathbf{c}}^1$ are aggregate consumption before and after the reform, respectively.
2. Consumption age-profile: $\mathbf{c}_j^1 = \frac{\mathbf{s}_{c,j}^1}{\bar{\mathbf{s}}_{c,j}^0}\mathbf{c}_j^0$, $\mathbf{h}^1 = \mathbf{h}^0$, where $\mathbf{c}_j^0$ and $\mathbf{c}_j^1$ denote consumption allocations at age j, and $\mathbf{s}_{c,j}^0$ and $\mathbf{s}_{c,j}^1$ are the consumption shares of age-j households before and after the reform, respectively.
3. Consumption distribution: the remaining consumption allocation changes after removing the level and age-profile changes from the total effects of the reform.

The changes in labor level, age-profile, and distribution are defined similarly. Note that the level and age-profile changes are straightforward to implement as they only require scaling the allocations, whereas the distributional change is more complex and is thus treated as the residual. For each change, we can compute the CEV based on the corresponding change in social welfare.

Table 13 presents the welfare decomposition results following the order of changes in Conesa et al. (2009). In particular, consumption allocation changes are introduced before labor allocation changes; and for both consumption and labor, the level change is added first, followed by the age-profile and the remaining distributional changes.[59]

In principle, the welfare decomposition results depend on the order in which changes are introduced in the thought experiment. To address this, we follow the Shapley-Owen-Shorrocks scheme as outlined in Audoly et al. (2025). In simple terms, we compute the weighted average of the decomposition results across all possible orders of changes. This approach ensures that the decomposition is order-independent and additively separable. More details can be found in Audoly et al. (2025). The corresponding results are presented in Table 4 in the main text and are very similar to those in Table 13.

---

[59] Conesa et al. (2009) do not separate the age-profile and the remaining distributional changes, and the combination of these two changes here corresponds to the distribution effects in their original paper.

Table 13: Alternative Decomposition of Welfare Gains

	Annual Tax	Lifetime Tax	
	Optimal	Status Quo μ	Optimal
Total Welfare Gain	0.19%	0.63%	1.27%
Consumption	−1.44%	0.53%	−1.90%
Level	−3.32%	0.38%	−5.24%
Age	0.35%	0.15%	0.74%
Distribution	1.59%	−0.01%	2.76%
Labor	1.65%	0.10%	3.24%
Level	1.39%	0.79%	3.89%
Age	0.06%	−0.12%	−0.12%
Distribution	0.20%	−0.57%	−0.52%

Notes: This table presents the welfare decomposition in the following order of changes: consumption level, age-profile, distribution, and then the corresponding labor changes.

C.3 Government Debt in Transition

In this section, we consider the dynamics of government debt after a permanent tax reform as described in Section 5.5. We show that if i) the new policy is only applicable to new households, ii) both the original and new tax policies satisfy the within-cohort government budget constraint, and iii) factor prices (i.e., interest rate and wage) are fixed, then the level of government debt will converge automatically to the sustainable level at the new stationary equilibrium. In other words, such tax reform is feasible.

Let B_t denote the level of government debt at the start of period t, then the period-t government budget constraint can be written as:

$$B_{t+1} = (1+r)B_t + G - X_t,$$

where X_t represents the net revenue of government in period t, G is government expenditures, and r is the interest rate. Differentiating both sides of the equation with respect to time t, we have

$$\Delta B_{t+1} = (1+r)\Delta B_t - \Delta X_t. \tag{28}$$

Suppose that the tax reform occurs at the beginning of period 0, and the new policy is only applied to new households, i.e., households born at $t \geq 0$. Let $\bar{x}_j$ and $\hat{x}_j$, $j = 1, \ldots, J$, denote the government's net revenue from age-j households under the original and new tax policies, respectively. Because a new cohort of households enter the economy in each period, and they live for J periods, we have

$$X_t = \begin{cases} \sum_{j=1}^{J} \bar{x}_j, & \text{if } t < 0, \\ \sum_{j=1}^{J} \hat{x}_j, & \text{if } t \geq J-1, \end{cases}$$

and

$$\Delta X_t = \begin{cases} 0, & \text{if } t < 0, \\ \hat{x}_{t+1} - \bar{x}_{t+1}, & \text{if } 0 \leq t \leq J-1, \\ 0, & \text{if } t > J-1. \end{cases}$$

Since the economy is at the original stationary equilibrium before period 0 with constant government debt, $\Delta B_0 = 0$. From equation (28), we then have

$$\begin{aligned} \Delta B_1 &= -\Delta X_0 = \bar{x}_1 - \hat{x}_1, \\ \Delta B_2 &= (1+r)(\bar{x}_1 - \hat{x}_1) + (\bar{x}_2 - \hat{x}_2), \\ &\vdots \\ \Delta B_J &= \sum_{j=1}^{J} (1+r)^{J-j} (\bar{x}_j - \hat{x}_j) \\ &= (1+r)^{J+1} \left[\left(\sum_{j=1}^{J} \frac{\bar{x}_j}{(1+r)^{j-1}} \right) - \left(\sum_{j=1}^{J} \frac{\hat{x}_j}{(1+r)^{j-1}} \right) \right]. \end{aligned}$$

Because both the original and new policies satisfy the same within-cohort budget constraint, we have

$$\sum_{j=1}^{J} \frac{\bar{x}_j}{(1+r)^{j-1}} = \sum_{j=1}^{J} \frac{\hat{x}_j}{(1+r)^{j-1}},$$

which then implies

$$\Delta B_J = 0,$$

and

$$\Delta B_t = (1+r)^{t-J} \Delta B_J = 0, \ \forall t \geq J+1.$$

That is, after the tax reform, although the level of government debt may vary in the short run, it will stabilize again starting from period J.

Furthermore, from the government budget constraint in period -1 and $J-1$,

$$\begin{aligned} B_0 &= (1+r)B_{-1} + G - X_{-1}, \\ B_J &= (1+r)B_{J-1} + G - X_{J-1}, \end{aligned}$$

we have

$$B_J - B_0 = (1+r)(B_{J-1} - B_{-1}) - (X_{J-1} - X_{-1}).$$

Because $\Delta B_J = 0$ and $\Delta B_0 = 0$, we have $B_J = B_{J-1}$ and $B_0 = B_{-1}$, the long-run level of government debt is then

$$B_J = B_0 + \frac{X_{J-1} - X_{-1}}{r} = B_0 + \frac{1}{r} \left[\sum_{j=1}^{J} (\hat{x}_j - \bar{x}_j) \right].$$

Note that this is exactly the level of government debt that balances the period government budget constraint at the new stationary equilibrium.

C.4 Basing the LIT on Average Earnings: Additional Results

Given the formula for $\widetilde{T}(y_j, Y, j)$ in the main text, summing taxes due over all working years yields

$$\sum_{j=1}^{J_R} \widetilde{T}(y_j, Y, j) = T\left(\sum_{j=1}^{J_R} y_j\right).$$

Therefore, under the alternative implementation of the LIT, lifetime tax liabilities are identical to those under the benchmark implementation as long as the T function remains the same. The only difference between the two implementations of the LIT lies in the *timing* of tax payments.

Figure 23 plots life-cycle profiles under the alternative implementation of the LIT, corresponding to Figure 4 for the benchmark implementation in the main text. The bottom-left panel shows that average after-tax earnings no longer decline sharply over the life cycle under the alternative implementation of the LIT, as the tax liability is now more evenly distributed across ages.

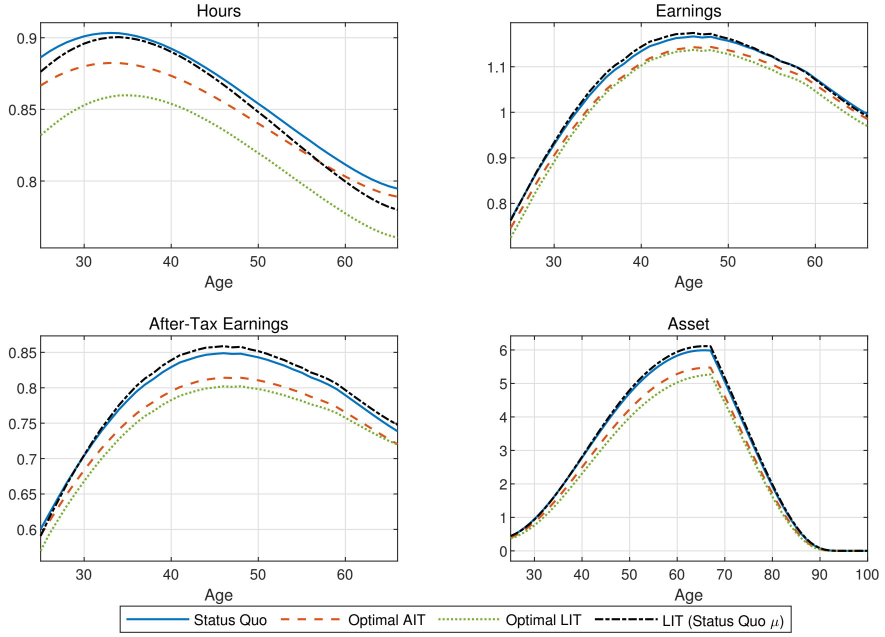

Figure 23: Life-Cycle Profiles of Household Averages
(LIT Based on Avg. Earnings)

Notes: This figure plots the life-cycle means in the status quo economy (blue solid lines), under the optimal annual income tax (red dashed lines), under the optimal lifetime income tax (green dotted lines), and under the lifetime income tax with the status quo tax progressivity (black dash-dotted lines). The lifetime income taxes are formulated based on average earnings.

C.5 Lump-Sum Transfers: Estimation of Tax Function

To estimate the HSV tax function with a lump-sum transfer, we follow the procedure in Wu (2021) to construct a mapping between household pre-government and post-government income. In brief, we use household income and transfer data from the Current Population Survey (CPS) 1999–2017 and compute tax liabilities with the NBER TAXSIM program. The three tax function parameters (μ, τ, ι) are then estimated via nonlinear least squares regression of post-government income on pre-government income, following Boar and Midrigan (2022). Consistent with the literature, introducing a lump-sum component reduces the estimated tax progressivity, with the estimated progressivity parameter μ becoming 0.093. The estimated lump-sum transfer ι is 0.072, corresponding to approximately \$3,920 in 2016 dollars.

C.6 Additional General Equilibrium Results

C.6.1 Life-Cycle Profiles in General Equilibrium

In this section, we present and discuss briefly the life-cycle implications of optimal annual and lifetime income tax reforms in general equilibrium. Figure 24 and Figure 25 display the life-cycle profiles in general equilibrium, corresponding to those in Figure 4 and Figure 6 for the baseline analysis.

As the general equilibrium effects lower the interest rate, households prefer to work less early in life, which leads to flatter life-cycle profiles of hours worked. Consequently, household earnings peak later around age 50. Under the optimal annual income tax, the lower early-life earnings and a more progressive tax policy significantly reduce household savings. In contrast, the optimal lifetime income tax redistributes income toward young households, raising their after-tax earnings and leading them to save more in anticipation of higher future taxes. It is worth noting that although total household savings increase under the optimal lifetime income tax and decrease under the optimal annual income tax, the net effect on physical capital remains similar between the two reforms. This is because the changes in government debt required to satisfy the stationary period government budget constraint absorb most of the difference in total household savings.

Labor efficiency improves over the life cycle under both optimal tax policies due to higher equilibrium wages, with greater efficiency gains under the lifetime income tax, consistent with our baseline findings. The life-cycle patterns of hours, earnings, and consumption inequality also largely resemble those in the baseline analysis.

C.6.2 Decomposition of Welfare Gains in General Equilibrium

To further understand the importance of general equilibrium, Table 14 reports step-by-step welfare changes from the status quo economy to the general equilibrium under the

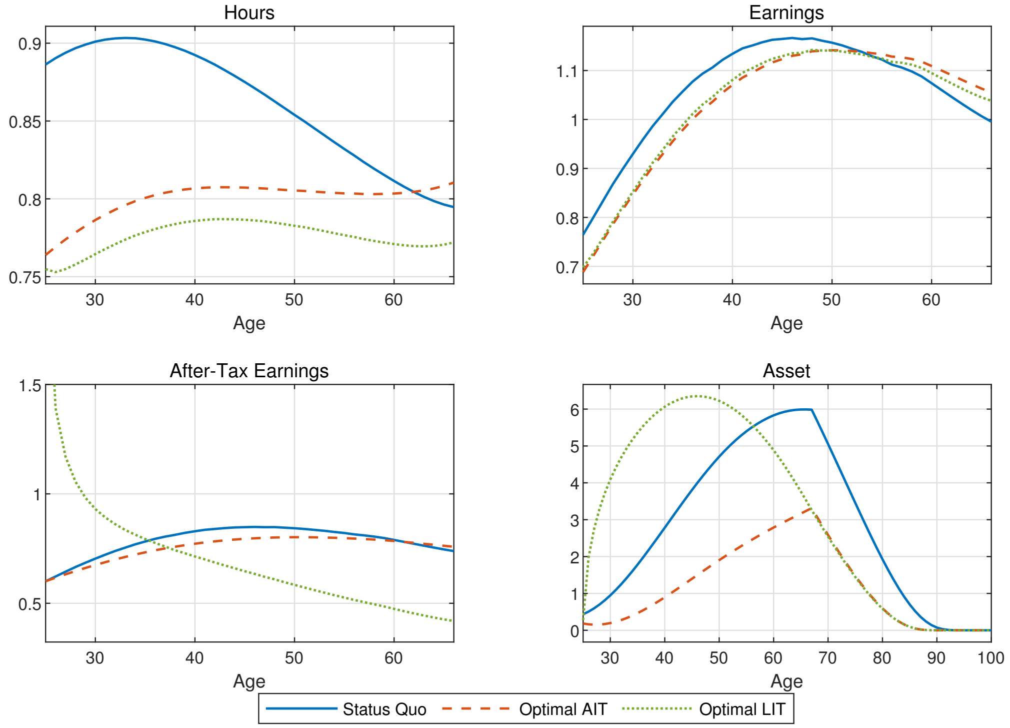

Figure 24: Life-Cycle Profiles of Household Averages
(General Equilibrium)

Notes: This figure plots the life-cycle profiles of cross-sectional means in the status quo economy (blue solid lines), under the optimal annual income tax (red dashed lines), and under the optimal lifetime income tax (green dotted lines).

optimal annual and lifetime income taxes. Consider first the optimal AIT (column "Annual Tax"). Starting from the status quo, we introduce the wage and interest rate changes reported in Panel B of Table 9 while maintaining the status quo tax policy. These changes capture the direct effects of factor price adjustments. The higher wage induces a welfare gain equivalent to 4.53% of lifetime consumption, whereas the lower interest rate leads to a welfare loss of 2.79%. Together (and accounting for their interaction), these factor price adjustments result in a net welfare gain of 1.55%.

Following these adjustments, however, the government budget constraint is no longer satisfied. Therefore, in a second step, we adjust the tax level to re-balance the government budget while keeping tax progressivity at the status quo. This captures the indirect effects of factor price adjustments through the government budget constraint. Since the higher wage and lower interest rate imply higher government revenues, the tax level can decline, generating additional welfare gains of 3.79%. Lastly, we implement the optimal AIT while holding the factor prices and tax revenue fixed (as in the partial equilibrium analysis), yielding the remaining welfare gain from tax reform, equivalent to 0.50% of lifetime consumption.

For the optimal LIT (column "Lifetime Tax"), the general equilibrium effects are similar

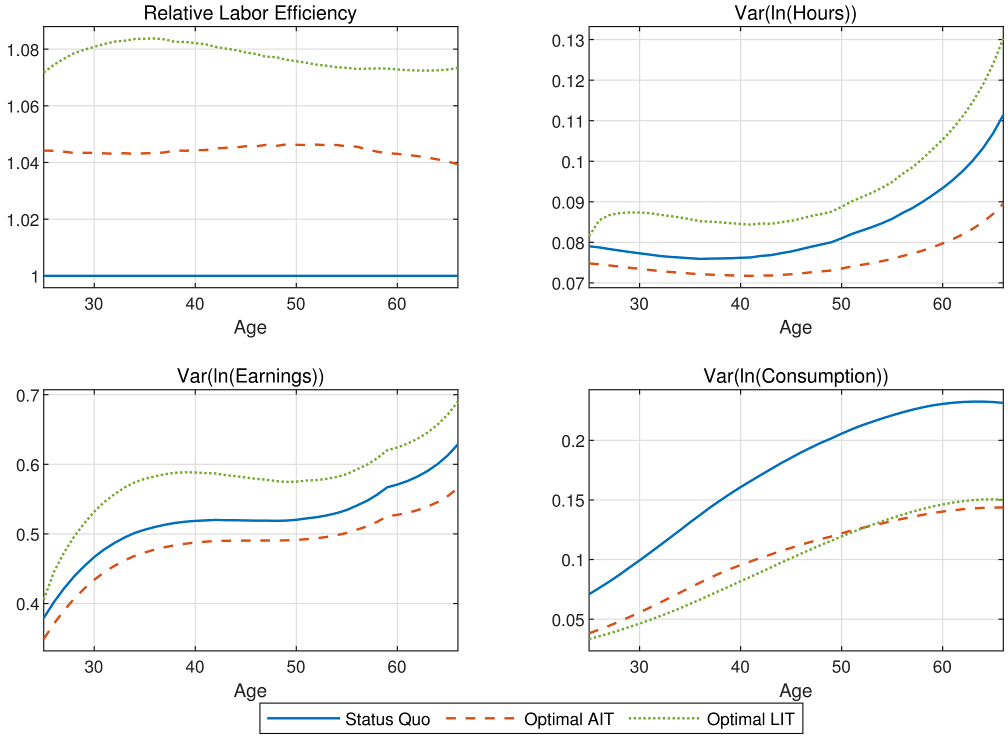

Figure 25: Life-Cycle Profiles of Labor Efficiency and Inequality (General Equilibrium)

Notes: This figure plots the life-cycle profiles of relative labor efficiency and cross-sectional variances in the status quo economy (blue solid lines), under the optimal annual income tax (red dashed lines), and under the optimal lifetime income tax (green dotted lines). Relative labor efficiency is labor efficiency divided by their values in the status quo economy.

Table 14: Welfare Gains: General vs. Partial Equilibrium Effects

	Annual Tax	Lifetime Tax
General Equilibrium Effects		
Factor Prices	1.55%	1.58%
Wage	4.53%	4.57%
Interest Rate	−2.79%	−2.81%
Interaction	−0.19%	−0.19%
Government Budget	3.79%	3.81%
Partial Equilibrium Effects		
Optimal Policy	0.50%	2.00%
Total Welfare Gain	5.84%	7.39%

Notes: The table reports step-by-step welfare changes from the status quo to the general equilibrium under optimal AIT or LIT.

to those under the optimal AIT, suggesting that the direct effects of factor price adjustments and their indirect effect through the government budget constraint are nearly identical. However, in the final step, switching to the optimal LIT yields an additional welfare gain of 2.00% of lifetime consumption, about 1.50% higher than that from the optimal AIT.

C.7 Within-Period Government Budget Constraint

To focus on the insurance and redistribution role of lifetime income tax, and avoid intergenerational transfers through the income tax system, our baseline analysis imposes a within-cohort government budget constraint when searching for the optimal policy. That is, the total revenue collected by the government from each cohort must remain constant when adjusting the income tax policy. In this section, we instead impose a within-period (across-cohort) government budget constraint such that the total revenue collected by the government in each period is constant across alternative polices.

For ease of comparison, Panel A of Table 15 reproduces the baseline results under the within-cohort government budget constraint. Panel B shows that once we switch to the within-period government budget constraint, for the annual income tax, the optimal policy becomes less progressive and achieves smaller welfare gains. In contrast, for the lifetime income tax, the opposite occurs, and the changes are more striking: the optimal tax progressivity increases from 0.242 to 0.582, and the welfare gain grows from 1.27% to 21.23% of lifetime consumption.

Table 15: Cohort vs. Period Government Budget Constraint

	Annual Tax		Lifetime Tax	
	Status Quo	Optimal	Status Quo μ	Optimal
A. Within-Cohort Budget Constraint				
Progressivity (μ)	0.137	0.194	0.137	0.242
Level (τ)	0.105	0.110	−0.431	−1.035
Level (comparable)	10.5%	11.0%	14.3%	17.7%
Avg. Tax Rate	13.8%	14.9%	15.7%	19.3%
Welfare Gain	–	0.19%	0.63%	1.27%
B. Within-Period (Across-Cohort) Budget Constraint				
Progressivity (μ)	0.137	0.154	0.137	0.582
Level (τ)	0.105	0.107	−0.509	−7.035
Level (comparable)	10.5%	10.7%	9.6%	8.6%
Avg. Tax Rate	13.8%	14.2%	11.0%	10.7%
Welfare Gain	–	0.02%	6.15%	21.23%

Notes: Welfare changes are in consumption equivalent variations as percentages of household lifetime consumption in the status quo. "Level (comparable)" is the average tax rate of a household with constant earnings equal to the status quo average in each working year. "Avg. Tax Rate" is the ratio between total labor income taxes and total labor earnings in the economy.

The exceptionally large welfare gain of the optimal lifetime income tax under the within-period budget constraint likely comes from intergenerational transfers from current to future households through the income tax system, which we verify by computing the transition path from the status quo economy to the optimal stationary equilibrium, similar to the exercises in Section 5.5. Government debt and expenditures are fixed at the status quo levels,[60] and hence the income tax level τ must adjust over time to balance the government period budget constraint along the transition, whereas the tax progressivity μ, and the income tax type if applicable, are changed once-and-for-all at time $t = 0$.

The top-left panel of Figure 26 presents the heterogeneous welfare effects of the annual income tax reform to different generations of households. Recall that generation x enters the economy at time $t = x$, and model age 1 corresponds to data age 25. On average, households born around the time of tax reform benefits more than older or future generations. The top-right panel shows that the tax level τ jumps up at the reform and then gradually increases to its level at the optimal stationary equilibrium.

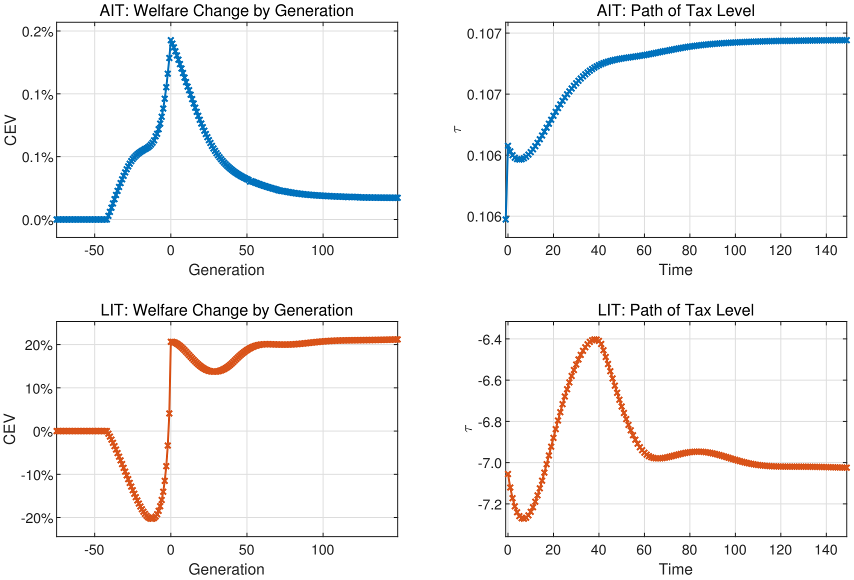

Figure 26: Transition to the Optimal Stationary Equilibrium (Within-Period Government Budget Constraint)

Notes: This figure shows the results related to the transition from the status quo economy to the optimal stationary equilibrium subject to the within-period government budget constraint with either annual (top panels) or lifetime (bottom panels) income taxes, under the assumption that the new policy is applied to all households. The left panels display the welfare effects of tax reform on different generations, and the right panels plot the transition path of income tax level τ that balances the government budget period-by-period. Generation x is born at time x, and tax reform occurs at time 0.

The bottom-left panel reveals that most of current working households suffer from the lifetime income tax reform, and the magnitudes of welfare losses are comparable to the large welfare gains enjoyed by future generations. The reason for the massive welfare losses is that most of current households did not receive the subsidies (i.e., negative taxes) to young

---

[60] Since the government now collects the same revenue per period in the original and optimal stationary equilibria, the level of government debt no longer needs to change.

households under the optimal lifetime income tax, but they still need to pay high taxes to finance such transfers to current and future young households. The bottom-right panel shows that the tax level τ drops to below its long-run optimal level at the commencement of transition; it then grows and overshoots before eventually falling back to the optimal stationary equilibrium level.