Time Averaging Meets Labor Supplies of Heckman, Lochner, and Taber

Sebastian Graves Victoria Gregory Lars Ljungqvist
Thomas J. Sargent*

May 29, 2023

Abstract

We incorporate time-averaging into the canonical model of Heckman, Lochner, and Taber (1998) (HLT) to study retirement decisions, government policies, and their interaction with the aggregate labor supply elasticity. The HLT model forced all agents to retire at age 65, while our model allows them to choose career lengths. A benchmark social security system puts all of our workers at corner solutions of their career-length choice problems and lets our model reproduce HLT model outcomes. But alternative tax and social security arrangements dislodge some agents from those corners, bringing associated changes in equilibrium prices and human capital accumulation decisions. A reform that links social security benefits to age but not to employment status eliminates the implicit tax on working beyond 65. High taxes with revenues returned lump-sum keep agents off corner solutions, raising the aggregate labor supply elasticity and threatening to bring about a "dual labor market" in which many people decide not to supply labor.

Key words: Time averaging, labor supply elasticity, retirement, taxation, Laffer curve, social security reform.

*Graves: Federal Reserve Board (email: sebastian.graves@nyu.edu); Gregory: Federal Reserve Bank of St. Louis (email:vgregory757@gmail.com); Ljungqvist: Stockholm School of Economics and New York University (email: lars.ljungqvist@hhs.se); Sargent: New York University and Hoover Institution (email: thomas.sargent@nyu.edu). Christopher Taber shared Heckman, Lochner, and Taber (1998a) computer codes that helped us replicate some results. We thank Marcos Sonnervig for excellent research assistance. Ljungqvist's research was supported by a grant from the Jan Wallander and Tom Hedelius Foundation. The views expressed here are solely those of the authors and do not necessarily reflect those of the Federal Reserve Bank of St. Louis, the Federal Reserve Board, or the Federal Reserve System.

1 Introduction

On the 25-year anniversary of the publication of the celebrated macro-labor analysis of Heckman, Lochner, and Taber (1998a, henceforth HLT), we extend their model to include additional determinants of labor supplies and study how they affect the aggregate labor supply issue, a controversial quantity in macroeconomics. HLT assumed an exogenous retirement age, a common assumption in macroeconomic applications of life-cycle models dating back at least to Auerbach and Kotlikoff's (1987) classic extension of the overlapping-generations structure of Diamond (1965) and Samuelson (1958) for quantitative policy analysis. HLT (p. 9) also assumed an inelastic labor supply before retirement and justified doing that by noting "[e]stimates of intertemporal substitution in labor supply estimated on annual data are small, so ignoring labor supply decisions will not greatly affect our analysis."

HLT's analysis represents the low-elasticity side of a long-standing debate in macroeconomics about the magnitude of the aggregate labor supply elasticity. Prescott represented the other side. In his 2004 Nobel lecture Prescott (2005, p. 385) deduced a high labor supply elasticity from observed employment fluctuations over the business cycle and an aggregation theory of Rogerson (1988) that he said "is every bit as important as the one giving rise to the aggregate production function." By combining (i) indivisible labor, with (ii) employment lotteries together with complete markets for insuring individual agents' consumption against the risk injected by those employment lotteries, Rogerson had deduced a high aggregate labor supply elasticity. By convexifying labor supply choices, those lotteries yield higher expected utilities and a large aggregate labor supply elasticity so long as the equilibrium employment-to-population ratio is less than one. Because they saw no micro evidence for them, critics of Rogerson's aggregation theory expressed doubts about components (ii) of the theory. Browning, Hansen, and Heckman (1999, p. 602) wrote that the "employment allocation mechanism strains credibility and is at odds with the micro evidence on individual employment histories."

But an alternative aggregation theory that also delivers a high labor supply elasticity discards Rogerson's employment lotteries and his assumption of complete consumption-insurance markets. This "time-averaging" theory replaces those components of Rogerson's theory with a life cycle model in which workers smooth across time and self-insure across random states by trading a risk-free one-period bond, thereby moving toward HLT's framework. Thus, in a continuous-time, non-stochastic life-cycle incomplete-market economy that retains indivisible labor component (i), Ljungqvist and Sargent (2006) deduced the same individual (expected) utilities, aggregate allocation and high aggregate labor supply elasticity that prevail in a Rogerson complete-market economy with employment lotteries. Instead of choosing probabilities of working at a point in time, each agent chooses a fraction of a lifetime to devote to work and uses a credit market to smooth consumption across episodes of work and times of not working called retirement. In a nutshell, Ljungqvist and Sargent showed that "time averaging" can replace lotteries and still deliver a high aggregate labor supply elasticity.[1]

When Prescott (2006a) discussed Ljungqvist and Sargent (2006) at the 2006 NBER macro conference, he welcomed a more plausible micro-foundation for a high aggregate labor supply elasticity and responded by adding a section "The Life Cycle and Labor Indivisibility" to what now appears as a second version (Prescott 2006b) of his original Nobel lecture (Prescott 2005) in which he extended the Ljungqvist and Sargent (2006) analysis to include a special intensive margin in agents' labor supply.

The shift from employment lotteries to time averaging between the two versions of Prescott's Nobel lecture frames new issues about possible sources of a high aggregate labor supply elasticity. In a Ljungqvist and Sargent time-averaging setting, a large aggregate labor supply elasticity no longer merely requires an employment-to-population ratio less than one; instead it requires workers to be at interior solutions for their choices of career lengths.[2]

Because advocates from both sides of the small-versus-large aggregate labor supply elasticity had adopted time-averaging as their aggregation theory, Ljungqvist and Sargent (2011) forecast a "labor supply elasticity accord" and described what that meant for interpreting more than half a century of cross-Atlantic discrepancies in labor market outcomes. This paper takes steps to promote such an accord by using a time-averaging version of HLT's model as our vehicle. Specifically, after including an operative extensive labor supply margin and a social security arrangement that approximates one that prevailed in the U.S. in the early 1970s, the time period that HLT used for their baseline steady state, we succeed in reproducing virtually all of HLT's analysis of skill-biased technological change. But in our model, because we have dropped HLT's exogenous retirement age, it is not hard-wiring but something else that makes the cousins of HLT's agents who live in our model retire at 65: *all*

[1] Independently, Chang and Kim (2006) discovered a high aggregate labor supply elasticity in simulations of a stochastic Bewley model with incomplete markets and indivisible labor. Their agents optimally alternate between periods of work and leisure (they "time average") to allocate consumption and leisure over their infinite lifespans.

[2] Note that while HLT adopt a life-cycle structure, their assumption of an exogenous retirement age puts their workers on a corner with respect to career length and thus turns off the Ljungqvist and Sargent source of high labor supply elasticity.

of our agents, who like HLT's, differ in their inherent abilities and their educational choices, nevertheless end up *choosing* to retire when they first become eligible for social security benefits. In our model, this behavior comes from everyone choosing corner solutions at kinks in their budget sets that are created by an implicit tax that the social security system imposed in the form of lost benefits that would be activated if they were to choose to work longer than a mandated age.[3]

After aligning outcomes in our model with HLT's, we conduct a first policy experiment by taking our time-averaging version of HLT's baseline economy and studying a social security reform that lets people receive social security benefits after the official retirement age, regardless of whether they choose to retire. This is our stylized version of an actual reform in U.S. social security that has actually taken effect only recently. That reform raises future benefits of agents who delay claiming social security enough so that they lose nothing in terms of the actuarial value of their lifetime benefits. By removing the kink in workers' budget constraints, the reform dislodges people from a corner at the official retirement age under the original social security arrangement. Our second policy experiment studies consequences of increases in the labor income tax rate above HLT's baseline setting. Prescott (2002) stressed that effects of labor tax rate increases on the aggregate labor supply depend on how a government spends associated increments in tax revenues: while government expenditures on very good substitutes for private consumption unleash a high labor supply elasticity, expenditures that are poor substitutes bring a low aggregate labor supply elasticity. In our time-averaging version of HLT's model, effects of tax rate increases also depend on whether workers are at corner solutions for career length at the official retirement age in our baseline economy; or whether, in our economy under the social security reform, they choose to work well into an interval of ages in which they experience steeply declining efficiency units of human capital (which would effectively be similar to retiring at a corner).

While our findings reaffirm some related analytical results in simple partial-equilibrium structures of Ljungqvist and Sargent (2006, 2014), we gather new insights from our model's ample HLT-style labor force heterogeneities, from dynamics contributed by our Ben-Porath human capital technologies, and from general equilibrium forces.

[3] As for our assumption that social security benefits not collected after age 65 cannot later be recouped, Schulz (2001, pp. 141-2) describes how this was the situation in the U.S. social security system between 1950 and 1972, after the repeal in 1950 of an earlier provision of a 1 percent increase in benefits for each year of delayed retirement. After 1972, a delayed retirement credit was reintroduced, but it is only with rules that recently became effective that the compensation is high enough for there to be no loss in the actuarial value of a worker's lifetime benefits.

Section 2 sets the stage by reproducing the quantitative analysis of HLT. Section 3 then describes our time-averaging version of the HLT framework. Section 4 analyzes a social security reform and Section 5 analyzes a tax-and-transfer policy. Section 6 analyzes a combination of both. Section 7 computes aggregate labor supply elasticities for these policy experiments. Section 8 conducts a sensitivity analysis with respect to new primitives at the core of our time-averaging version of HLT. It also addresses technical issues about computing equilibria as well as economically substantive issues associated with lumpiness in labor supply. Section 9 summarizes challenges that our findings pose for economic policy making.

2 The HLT Model

HLT formulated an overlapping generations model in discrete time at an annual frequency. Each person lives inside the model from age 18 to age 80. People who work supply labor inelastically until age 65. The cohort of 80-year olds that exits the model each year is replaced by an equal-sized new cohort of 18-year olds who come into the model with high school degrees and who either start to work immediately as high school graduates or choose to attend four-year college and then work as college graduates. An agent orders streams of a single consumption good by standard time-additively preferences with subjective discount factor δ. The utility from consuming C within a period is

$$U(C) = \frac{C^{\gamma} - 1}{\gamma}, \quad \gamma \geq 0, \tag{1}$$

so that marginal utility diminishes in C when $\gamma < 1$. By virtue of L'Hôpital's rule, utility function (1) becomes a logarithm as γ approaches zero. The elasticity of intertemporal substitution is constant at $(1-\gamma)^{-1}$.

Agents are heterogeneous with respect to their innate abilities and costs of attending college. When a cohort enters the economy with high school degrees at age 18, it is divided into four equal-sized ability groups indexed by $\theta \in \{1, 2, 3, 4\}$. All members of the same ability group share identical endowments of human capital, technologies for on-the-job human capital accumulation, and a probability distribution that governs an idiosyncratic nonpecuniary cost ϵ of attending college. The random cost ϵ is drawn from a normal distribution with standard deviation σ and ability-specific mean μ_θ. Given an ability θ and a realization ϵ of cost, an agent makes an irrevocable decision either immediately to work as a high school graduate (this is schooling choice $S = 1$) or to attend four-year college at an annual tuition cost of ζ and after graduating to work as a college graduate (this is schooling choice $S = 2$).

An agent of ability θ who makes schooling choice S has an initial endowment $H^S(\theta)$ units of S-specific human capital. While on the job, the agent can augment his human capital by diverting time from working to investing in human capital, with total per-period available time normalized to one. Thus, given a human capital stock H_n^S at age n, an agent who made schooling choice S can use a fraction I of his time to acquire age $n+1$ human capital according to a Ben-Porath technology

$$H_{n+1}^S = A^S(\theta)\, I^{\alpha_S} \left(H_n^S\right)^{\beta_S} + H_n^S \,, \tag{2}$$

where $A^S(\theta) > 0$, $0 < \alpha_S < 1$ and $0 \leq \beta_S \leq 1$.[4] Hence, the ability-specific talents of an agent in ability group θ are encoded in an initial human capital endowment $H^S(\theta)$ and a multiplicative factor $A^S(\theta)$ of the human capital technology. There is a pair $\big(H^S(\theta), A^S(\theta)\big)$ for each schooling choice $S = 1, 2$. Parameters α_S and β_S are common to high school graduates ($S = 1$) or college graduates ($S = 2$), respectively.

The aggregate production function exhibits constant returns to scale

$$F(\bar{H}^1, \bar{H}^2, \bar{K}) = a_3 \left((1 - a_2) \left[a_1 \left(\bar{H}^1\right)^{\rho_1} + (1 - a_1) \left(\bar{H}^2\right)^{\rho_1} \right]^{\rho_2/\rho_1} + a_2 \bar{K}^{\rho_2} \right)^{1/\rho_2} . \tag{3}$$

Inputs are the aggregate physical capital stock $\bar{K}$ and the aggregate human capital $\bar{H}^S$ supplied by high school ($S = 1$) and college ($S = 2$) workers, respectively. Notice that the right side of equation (3) excludes time and therefore also human capital that is used to create human capital next period. Elasticities of substitution are $(1-\rho_1)^{-1}$ between high school and college human capital and $(1-\rho_2)^{-1}$ between physical capital and composite human capital, respectively. Output can be allocated to private consumption, government consumption, or additional physical capital for next period.[5] Government consumption G enters neither agents' preferences nor the technologies for producing goods and human capital.

Since there are no aggregate shocks and all individual uncertainty is resolved when agents enter the economy, competitive equilibria are computed under perfect foresight. Markets are

[4] HLT (p. 19) assume no depreciation of human capital so as to be "consistent with the lack of any peak in life-cycle wage–age profiles reported in the literature."

[5] HLT assume no depreciation of physical capital. However, if we were to want to accommodate reductions in the capital stock within an equilibrium, we could simply adopt an assumption of reversible capital, i.e., by assuming that physical capital can one-for-one be converted into goods for consumption.

assumed to be complete. Agents can freely lend and borrow at a risk-free interest rate.[6] Equilibrium net savings equal the stock of physical capital. A government policy consists of flat-rate taxes τ_l and τ_k on households' labor income and capital income, respectively. The tax rate on capital income is also the rate at which interest expenses on borrowings are deductible from an agent's tax liability. Hence, borrowers and lenders both face an effective interest rate equal to the market interest rate multiplied by $(1-\tau_k)$. Net tax revenues are all used to finance government consumption. As explained in the next subsection, HLT adopted an auxiliary assumption of a lump-sum transfer from 65-year olds to 18-year olds in order to target a capital-output ratio.

2.1 Calibration and estimation

HLT used both micro- and macroeonomic data to calibrate parameters. They wanted their model of human capital and earnings dynamics to match education and earnings outcomes of white male civilians using National Longitudinal Survey of Youth (NLSY) data for the period 1979–1993. To infer parameters of the aggregate production function they use aggregates from the National Income and Product accounts as well as data on workers in the Current Population Survey (CPS) for the period 1963–1993. Tables 1, 2, and 4 summarize parameters inferred by HLT.

Following a common calibration approach in applied macroeconomics, HLT (p. 15) posit "discount [$\delta = 0.96$] and intertemporal substitution [$\gamma = 0.1$] parameters in consumption to be consistent with those reported in the empirical literature and that enable us to reproduce key features of the macro data – like the capital–output ratio." Besides a capital–output ratio of 4, HLT seek to target a steady state after-tax interest rate of $(1-\tau_k)r = 0.05$, to which we return below. HLT (pp. 26-27) assume a uniform tax rate on labor and capital income, $\tau_l = \tau_k = 0.15$, "that was suggested by Pechman (1987) as an accurate approximation to the true rate over our sample period once itemizations, deductions, and income-contingent benefits are factored in." Not reported by HLT, but gathered from Taber (2002, Table 1), annual college tuition ζ is set equal to 1.02 (thousands of 1992 dollars).

Using data from the NLSY and measures of college tuition collected by the Department

[6] HLT (p. 10) note the absence of "short-run credit constraints that are often featured in the literature on schooling and human capital accumulation. Our model is consistent with the evidence presented in [e.g., Cameron and Heckman (1998)] that long-run family factors correlated with income (the θ operating through $A^S(\theta)$ and $H^S(\theta)$) affect schooling, but that short-term credit constraints are not empirically important. . . . The mechanism generating the family income–schooling relationship operates through family-acquired human capital and not credit rationing."

Table 1: Same or similar parameterizations across HLT's and our models

Parameter	HLT	Our*	Description
$\underline{\eta}$	18	=	age entering economy
$\overline{\eta}$	80	=	age exiting economy
δ	0.96	=	discount factor
γ	0.1	0	intertemporal substitution
τ_l	0.15	=	labor income tax rate
τ_k	0.15	=	capital income tax rate
ζ	1.02	=	annual tuition in 4-year college
			Ben-Porath [high school / college]
α_1 / α_2	0.945 / 0.939	=	exponent on investment
β_1 / β_2	0.832 / 0.871	=	exponent on human capital
$A^1(1)$ / $A^2(1)$	0.081 / 0.081	=	productivity, ability $\theta = 1$
$A^1(2)$ / $A^2(2)$	0.085 / 0.082	=	$\theta = 2$
$A^1(3)$ / $A^2(3)$	0.087 / 0.082	=	$\theta = 3$
$A^1(4)$ / $A^2(4)$	0.086 / 0.084	=	$\theta = 4$
$H^1(1)$ / $H^2(1)$	8.042 / 11.117	=	initial human capital, $\theta = 1$
$H^1(2)$ / $H^2(2)$	10.063 / 12.271	=	$\theta = 2$
$H^1(3)$ / $H^2(3)$	11.127 / 12.960	=	$\theta = 3$
$H^1(4)$ / $H^2(4)$	10.361 / 15.095	=	$\theta = 4$
			Production function for goods
a_1	0.496†	0.475	share high school (versus college)
a_2	0.252†	0.235	share physical capital (versus human)
a_3	2.504†	2.554	productivity
			substitution between
ρ_1	0.306	=	high school and college
ρ_2	-0.034	0	human and physical capital

* An equality sign means that our model adopts HLT's parameter value.

† HLT report no parameter values for a_1, a_2, and a_3, which we instead deduce from our replication of HLT in Section 2.2.

of Education, HLT estimated heterogeneities in human capital endowments, human capital production technologies, and propensities to attend college. First, they constructed four equal-sized ability groups $\theta \in \{1, 2, 3, 4\}$ from scores on 1980 Armed Forces Qualifying Test (AFQT) administered by the NLSY. Second, for each ability group θ and schooling choice S, they used data on earnings to estimate human capital endowment $H^S(\theta)$ and the parameters $A^S(\theta)$, α_S, and β_S in human capital technology (2). For any particular set of those parameters and under HLT's premise "that interest rates and the after-tax rental rates on human capital are fixed at constant but empirically concordant values," earnings profiles in the model can be deduced after computing optimal human capital investments over the lifecycle. Then, by using nonlinear least squares estimation to minimize the discrepancy between those deduced earnings profiles and actual ones in the NLSY, HLT obtained the estimates of human capital parameters in Table 1. Both $H^S(\theta)$ and $A^S(\theta)$ are mostly increasing in ability θ except that high school graduates of ability group 3 have larger human capital endowment $H^1(3)$ and higher productivity $A^1(3)$ in the human capital technology when compared to ability group 4, differences consistent with earnings profiles for the two groups in the NLSY.

Third, HLT estimated an ability-specific stochastic process for the nonpecuniary cost ϵ of attending college. Under the assumptions of complete markets and inelastic labor supplies, utility-optimizing human capital investments maximize present values of lifetime after-tax labor earnings. HLT used maximization of the present value of after-tax labor earnings to characterize schooling decisions. Thus, an agent chooses to attend college if the present-value after-tax labor earnings as a college graduate, reduced by tuition and the nonpecuniary cost ϵ of attending college, exceeds what the present value of after-tax income he would have earned as a high school graduate. The shock ϵ is expressed in dollars; its ability-specific means μ_θ are reported in Table 2, where negative numbers represent costs and positive numbers indicate a nonpecuniary benefit of attending college. For the first three ability groups, the mean cost decreases as ability rises, eventually turning into a mean benefit for ability group 3. However, this changes for the highest ability group 4 whose mean cost is the second largest, only exceeded by the mean cost of the lowest ability group 1. The common standard deviation σ of the ϵ-shocks is identified and estimated based on differences in college tuition across U.S. states.

To infer parameters of the aggregate production function, HLT started with the usual decomposition of aggregate output into compensation to labor and capital based on the National Income and Product accounts. They inferred two human capital aggregates, high school human capital $\bar{H}^1$ and college human capital $\bar{H}^2$ together with corresponding skill

Table 2: Nonpecuniary college cost $\epsilon \sim N(\mu_\theta, \sigma)$, expressed in thousands of dollars in HLT versus utils in our model*

Parameter	HLT	Our	Description
μ_1	-53.02	-6.819	ability-specific mean, $\theta = 1$
μ_2	-2.82	-2.808	$\theta = 2$
μ_3	29.77	-0.912	$\theta = 3$
μ_4	-28.65	-4.587	$\theta = 4$
σ	22.41	1.5	standard deviation

* A negative (positive) number represents a nonpecuniary cost (benefit) of attending college.

prices, R^1 and R^2 from CPS data. They constructed the stock of physical capital $\bar{K}$ from Federal Reserve Board data sets. Except for the share parameters on the two human capital aggregates, they assumed that other parameters of production function (3) are time invariant; HLT specified a linear time trend for $\log[(1-a_1)/a_1]$, and inferred a trend coefficient $\varphi = 0.036$. While such evidence on skill biased technological change informed estimates of parameters of production function (3), HLT did not use it to calibrate other parameters.[7]

When they set $\varphi = 0$, with one additional assumption the remaining parameters in Tables 1, 2, and 4 generated HLT's initial steady before the mid 1970. To attain the calibration target of a capital-output ratio equal to 4, HLT (p. 27) imposed transfers from each retiring cohort to each new cohort that enters the labor market. Thus, "for each year, transfer X is taken from all workers at retirement age [65], and the total amount is equally distributed to all individuals (irrespective of ability) of age [18] in that period. For the simulations reported in this paper, $X \approx \$30,000$."

In their main analysis, HLT (p. 28) "consider a permanent shift in technology toward skilled labor .... We start from an initial steady state [$\varphi = 0$] and suppose that the technology begins to manifest a skill bias in the mid 1970s [$\varphi = 0.036$] ... and continuing for 30 years" after which the economy converges to a new steady state. The shift in technology

[7] While acknowledging that the estimation of the human capital production technology "ignores the price variation induced by technological change," HLT (p. 19) point out that "a remarkable finding of our research, reported in [HLT's] Appendix B, is that this misspecification of the economic environment has only slight consequences for the estimation of the curvature parameters of the human capital technology."

comes as a complete surprise to the agents. But once the technology change begins, agents know the entire future path of technology. Hence, from the mid 1970s HLT computed a perfect foresight equilibrium. In Table 3 we report some aspects of the initial and the future steady states.[8]

2.2 Replication of HLT

Because HLT reported neither a tuition cost ζ nor their estimates of production function parameters $\{a_1, a_2, a_3\}$, we must do two more things in order to replicate HLT's findings. First, by targeting prices in HLT's baseline steady state, we can infer values of the three unreported production parameters. Second, after adding those three parameter values and a tuition cost from Taber (2002, Table 1) to the parameterization reported in HLT, we can reproduce outcomes in HLT's baseline steady state, as well as outcomes in HLT's second steady state in our subsequent simulation of their experiment of skill-biased technological change.

The first two columns in Table 3 report our reconstructions of those two HLT steady states. Because the above calibration of the three production parameters targeted prices $\{r, R^1, R^2\}$ in HLT's baseline steady, these prices are replicated perfectly. In addition, our computed prices in the second steady state are very close to HLT's reported numbers $\{0.0609, 2.23, 2.41\}$ (see the text inside HLT's Figures 5 and 8). Likewise, our replicated end-of-life human capitals for different ability groups and schooling levels in the baseline steady state are very similar to those in HLT's Table I. A visual inspection of the leftmost points in HLT's Figure 9 shows that their "utilized" aggregates of human capital in the baseline steady state are close to our numbers $\bar{H}^1 = 274$ and $\bar{H}^2 = 280$. HLT do not report counterparts of the college enrollment rates that we report in our Table 3. Our Appendix A.1, Figure A.1 further confirms that we do a good job of reproducing HLT's quantities, including lifecycle earnings profiles and human capital investments on the job.[9]

[8] Regarding rental rates on human capital in the baseline steady state, HLT (p. 27) "calibrate the aggregate production parameters to yield . . . rental rates on human capital of 2. These values are consistent with those used in estimation of the human capital production parameters. Since human capital is measured in terms of hourly wages, earnings from our simulations are annual income measured in thousand of dollars if agents work 2000 hours per year."

[9] An unresolved discrepancy between HLT and our replication is that our present-value earnings of high-school graduates of different abilities are between 6.9-7.0 percent lower than those reported in HLT's Table II, and for college graduates 7.1-7.2 percent lower.

Table 3: Steady states in HLT and our model

Variable	HLT*		Our model	
	Baseline	SBTC†	Baseline	SBTC†
Retirement age, all workers	65	65	65	65
Interest rate, r	0.0588	0.0609	0.0588	0.0599
Rental rate on human capital				
high school, R^1	2	2.20	2	2.27
college, R^2	2	2.39	2	2.45
Inputs in goods production				
high school, $\bar{H}^1$	274	119	249	94
college, $\bar{H}^2$	280	446	287	459
physical capital, $\bar{K}$	5725	6814	5605	6849
College enrollment rates‡				
ability $\theta = 1$	0.09	0.38	0.11	0.47
$\theta = 2$	0.28	0.67	0.34	0.77
$\theta = 3$	0.56	0.89	0.56	0.90
$\theta = 4$	0.81	0.99	0.86	0.99
End-of-life human capital for high school / college workers				
ability $\theta = 1$	9.4 / 13.5	9.0 /12.9	9.2 / 13.0	9.0 / 12.8
$\theta = 2$	12.1 / 14.9	11.5 / 14.2	11.8 / 14.4	11.5 / 14.1
$\theta = 3$	13.6 / 15.5	12.9 / 14.8	13.2 / 15.0	12.8 / 14.7
$\theta = 4$	12.6 / 18.2	12.0 / 17.4	12.3 / 17.6	12.0 / 17.2

* Outcomes in HLT are based on our reconstruction of HLT.

† SBTC is the steady state after Skill-Biased Technological Change.

‡ College enrollment rates in the baseline steady state in HLT are based on our reconstruction of HLT, whereas the corresponding numbers in our model are those of Taber (2002, Table 1), to which our model is calibrated.

3 Our model of time averaging

We want to endogenize agents' choices of career lengths while retaining as much as possible of HLT's framework. So we introduce no intensive margin of labor supply, only an extensive margin that involves choosing career length under the maintained assumption that an agent incurs an additively separable disutility B of work each period (as well as during college attendance). To mimic HLT's hard-wired outcome that all agents retire at age 65 regardless of their abilities or schooling choices, we introduce a social security program with special features designed to align with the U.S. system. In that way, we endogenize the outcome that agents all retire at age 65, something that had been hard-wired by many authors besides HLT. Actually, for a long time the U.S. had prescribed that agents who work beyond an official retirement age would lose the value of social security benefits that they would collected if they had instead retired. Hence, any labor income earned after age 65 was subject to an extra implicit tax that could explain why multiple types of agents would optimally *choose* to retire at the official retirement age. We note here that to rule out excessively large labor supply responses in old age in some of our subsequent policy experiments, we assume that efficiency units of human capital depreciate with age. Consistent with HLT's assumptions, such depreciation is first noticeable only when agents reach their 60s.

In addition to these extensions of HLT's model, we make two auxiliary modifications. First, while retaining HLT's specification for an idiosyncratic nonpecuniary random cost of attending college, we express the shock in utils rather than dollars. This is natural for our model of endogenous career length because utility maximization is then no longer tantamount to maximizing the present value of after-tax labor income, as it was in HLT's model. Second, our introduction of a pay-as-you-go social security program reduces aggregate savings, so to attain the targeted capital-output ratio we can't rely on HLT's auxiliary assumption of administering a lump-sum transfer from 65-year olds to 18-year olds. Instead, we assume that agents who live inside our model hold only a fraction of the equilibrium capital stock and that the rest is held by 'investors' whom we do not explicitly model, an approach that retains interest rate endogeneity in model simulations.[10] Another justification for this modification

[10]Because plain vanilla life-cycle models typically fail to explain observed high wealth inequalities and fall short of explaining levels of aggregate savings as measured by an economy's stock of physical capital, researches have activated other features that affect wealth accumulation, including entrepreneurship and bequest motives (see e.g. De Nardi (2015) for a survey of the literature). Our assumption that agents who live inside the model hold only a fraction of the equilibrium capital stock acknowledges such features omitted from our framework. In contrast to earlier studies that use such an assumption to deduce a fixed interest rate at which the analyses are conducted (see e.g. Storesletten, Telmer, and Yaron (2004)), we keep the fraction

as we switch from HLT's inelastic labor supply to our model of endogenous career length is that, especially for low-ability agents with meager labor earnings prospects, lump-sum transfers become a potent determinant of labor supplies. By doing away with HLT's lump-sum transfers, we ensure that an auxiliary assumption for targeting a capital-output ratio does not unduly affect other equilibrium outcomes.

Finally, HLT's values for the preference parameter γ and the technology parameter ρ_2 are both close to zero. We set those parameters to zero. Thus, instead of HLT's utility function (1), for us utility from consuming C and either working ($\omega = 1$) or not working ($\omega = 0$) in a period is given by

$$\tilde{U}(C, N) = \log(C) - B\,\omega\,, \tag{4}$$

where $\omega = 1$ also during college attendance. When combined with an additively separable disutility B of working, the logarithm of consumption specification makes our preferences consistent with balanced growth. Taking a limit as parameter ρ_2 approaches zero, we replace HLT's production function (3) by

$$\tilde{F}(\bar{H}^1, \bar{H}^2, \bar{K}) = a_3 \left[a_1 \left(\bar{H}^1\right)^{\rho_1} + (1 - a_1) \left(\bar{H}^2\right)^{\rho_1}\right]^{(1-a_2)/\rho_1} \bar{K}^{a_2}, \tag{5}$$

which is a Cobb-Douglas production function in the physical capital stock and a CES aggregator for human capital having share parameters a_2 and $(1 - a_2)$, respectively.

In addition to the disutility B of working that appears in utility function (4), other new parameters in our model are listed in Table 4. To describe how human capital gets converted into age-dependent efficiency units, let

$$e(n) = \frac{1}{1 + \exp(\phi_1(n - \phi_2))} \leq 1 \tag{6}$$

be a multplicative factor that translates the human capital stock of an agent of age n into efficiency units. Thus, given a schooling choice S, an agent of age n with human capital stock H^S has $e(n)H^S$ efficiency units that can be employed in the goods production technology (5). Note that HLT's human capital technology (2) remains unaltered. Our baseline social security program consists of a payroll tax rate τ_p levied on all labor income, plus a social security benefit P paid to all agents of age $\eta_p = 65$ or older who are not working. The benefit P is taxed at the labor income tax rate τ_l. Agents who live inside the model are assumed to

of capital held by agents in the model fixed across policy experiments and consequently have an endogenous interest rate.

hold a share κ of the aggregate capital stock; the remaining share $1-\kappa$ is held by 'investors' who live outside the model but whose capital income is also taxed by the government.

Following HLT, government consumption G is set endogenously by requiring that the government budget constraint, including our social security program, is satisfied.

Table 4: Model-specific features

Parameter	**Value**	**Description**
		HLT
η_R	65	exogenous retirement age
X	30	lump-sum transfer from old to young
		Our model
B	0.8	disutility of working
τ_p	0.1	payroll tax rate
η_p	65	age of eligibility for social security
P	8	social security benefit
κ	0.388	fraction of capital held by agents
		efficiency units of human capital
ϕ_1	0.2	slope coefficient
ϕ_2	75	age at inflection point

3.1 Choices

Agents choose a schooling level first, then a retirement age, and then paths for consumption, saving, and human capital investment. Let's work backwards.

Conditional on retiring at age $\hat{n}$, the value function of an employed agent of age $n < \hat{n}$, type θ, schooling level S, savings K, and human capital H satisfies the Bellman equation

$$V_n^{\hat{n}}(H, K, S, \theta) = \max_{C,I,K',H'} \Big[\log(C) - B + \delta V_{n+1}^{\hat{n}}(H', K', S, \theta)\Big] \tag{7}$$

subject to

$$C + K' = (1 + (1 - \tau_k)r)K + (1 - \tau_l - \tau_p)R^S e(n)(1 - I)H, \tag{8}$$

$$H' = H + A^S(\theta) I^{\alpha_S} H^{\beta_S}. \tag{9}$$

The value function for this same type of agent during retirement, $n \geq \hat{n}$, satisfies

$$V_n^{\hat{n}}(H, K, S, \theta) = \max_{C, K'} \left[\log(C) + \delta V_{n+1}^{\hat{n}}(H, K', S, \theta) \right] \tag{10}$$

subject to

$$C + K' = (1 + (1 - \tau)r)K + (1 - \tau_l)\mathbb{1}(n \geq \eta_p)P, \tag{11}$$

where $\mathbb{1}(n \geq \eta_p)$ is an indicator function equal to 1 if $n \geq \eta_p$, and zero otherwise. In the final year $n = \overline{\eta}$, the value function on the right side of the Bellman equation is zero, $V_{\overline{\eta}+1}^{\hat{n}}(\cdot) = 0$, and maximization is subject to an additional constraint requiring the worker to be solvent, $K_{\overline{\eta}+1} \geq 0$.

During the first four years of life, an $S = 2$ agent who attends college confronts a consumption-saving decision. Specifically, at ages $n = \underline{\eta} + j$ for $j = 0, 1, 2, 3$, the agent's optimization problem (7)–(9) is modified as follows. Budget constraint (8) is altered to become

$$C + K' + \zeta = (1 + (1 - \tau_k)r)K, \tag{12}$$

where college tuition ζ is added to the left side, and there is no labor income on the right side. Furthermore, we impose $I = 0$ during college attendance so that by law of motion (9), the agent's college human capital by graduation is equal to the endowment, $H^2_{\underline{\eta}+4} = H^2(\theta)$. Instead of earning a college degree, the agent could have started working immediately at age $\underline{\eta}$ as a high school graduate with high school human capital equal to the endowment, $H^1_{\underline{\eta}} = H^1(\theta)$. Every agent enters the economy with zero savings, $K_{\underline{\eta}} = 0$. Conditional on a retirement age $\hat{n}$, we can use a shooting algorithm to solve this part of the problem.

Conditional on a schooling choice S, an optimal retirement age $\hat{n}$ solves

$$\hat{V}^S(\theta) = \max_{\hat{n}} \left\{ V_{\underline{\eta}}^{\hat{n}}(H^S(\theta), 0, S, \theta) : \ \hat{n} = \underline{\eta} + 4(S - 1), \ldots, \overline{\eta} + 1 \right\}, \tag{13}$$

where the choice $\hat{n} = \overline{\eta} + 1$ means that the agent never retires. Following HLT, a decision to attend college ($S = 2$) is a four-year commitment which, in our analysis, means that the agent cannot choose to retire until after graduation, $\hat{n} \geq \underline{\eta} + 4$.[11] Finally, an agent who has

[11] Like HLT, we assume a nonpecuniary random cost ϵ of attending college for four years; it is a cost if the realization of ϵ is negative, a benefit if the ϵ realization is positive.

drawn a nonpecuniary college cost ϵ attends college if

$$\hat{V}^2(\theta) + \epsilon \geq \hat{V}^1(\theta), \tag{14}$$

where $\epsilon \sim N(\mu_\theta, \sigma)$. We let $p(S|\theta)$ denote the probability that an agent of ability θ makes schooling choice S.

Throughout the above description of an agent's optimization problem, we have proceeded under the assumption that it is optimal for an agent to front-load, i.e., to work during a first part of life and to enjoy retirement during a second part. Five features ensure the optimality of such a career strategy. First, as in HLT, the calibration of our time averaging model is such that the equilibrium after-tax interest rate is greater than the subjective rate of discounting, $(1-\tau_k)r > \delta^{-1} - 1$, and hence, it is optimal for an agent to generate labor income early in life in order to earn the high market interest rate on any lifecycle savings. Second, after learning his ability type θ and idiosyncratic nonpecuniary cost ϵ of attending college, an agent faces no further uncertainty and therefore, the optimal choice of lifetime labor supply is deterministic and requires no contingency plans to adjust future labor supplies to shocks. Third, the assumption that efficiency units of human capital depreciate with age means that a postponement of labor supply is associated with a less advantageous ratio of efficiency units to human capital. Fourth, the disadvantage of postponing labor supply is further accentuated under the social security system in our baseline economy in which all labor income after age 65 is subject to an extra implicit tax in the form of lost social security benefits that could have been collected if a worker had instead retired. Fifth, under HLT's and our common assumption that borrowing and lending are risk-free, an agent can choose any lifetime consumption profile consistent with his present-value budget constraint; for example, an agent can mortgage his future social security benefits. That could be a good decision if lifetime labor supply calls for early retirement before the official retirement age.

3.2 Profit maximization

Competitive firms operate the production technology shown in equation (5). The following first-order conditions determine skill prices and the return on capital:

$$R^1 = a_3(1-a_2)a_1 Q^{1-a_2-\rho_1}\bar{K}^{a_2}(\bar{H}^1)^{\rho_1-1} \tag{15}$$

$$R^2 = a_3(1-a_2)(1-a_1) Q^{1-a_2-\rho_1}\bar{K}^{a_2}(\bar{H}^2)^{\rho_1-1} \tag{16}$$

$$r = a_3 a_2 Q^{1-a_2}\bar{K}^{a_2-1} \tag{17}$$

where Q is a measure of total labor supplied to firms

$$Q = \left[a_1(\bar{H}^1)^{\rho_1} + (1-a_1)(\bar{H}^2)^{\rho_1}\right]^{\frac{1}{\rho_1}}. \tag{18}$$

3.3 Stationary equilibrium

A parameterization is said to be 'regular' if it supports a stationary equilibrium that satisfies two conditions: (i) all agents optimally front-load their labor supply, and (ii) conditional on a schooling choice, solutions to all agents' optimization problems are unique. We call such an equilibrium a *regular stationary equilibrium*.

Definition: A *regular stationary equilibrium* is a per-person allocation of consumption $\hat{C}_n(S,\theta)$, human capital $\hat{H}_n(S,\theta)$, fraction of time $\hat{I}_n(S,\theta)$ devoted to on-the-job human capital accumulation, and savings $\hat{K}_{n+1}(S,\theta)$, indexed by age $n = \underline{\eta}, \underline{\eta}+1, \underline{\eta}+2, \ldots, \overline{\eta}$, schooling $S = 1,2$, and ability $\theta = 1,2,3,4$, with associated retirement age $\hat{n}(S,\theta)$, and probability $p(S|\theta)$ that an agent of ability θ makes schooling choice S; aggregate quantities $\{\bar{K}, \bar{H}^1, \bar{H}^2\}$ of inputs in goods production; prices $\{r, R^1, R^2\}$; and government policy $\{\tau_l, \tau_k, \tau_p, \eta_p, P, G\}$ such that:

1. Given prices $\{r, R^1, R^2\}$ and government policy $\{\tau_l, \tau_k, \tau_p, \eta_p, P\}$, for each pair (θ, S) the allocation $\{\hat{C}_n(S,\theta), \hat{H}_n(S,\theta), \hat{I}_n(S,\theta), \hat{K}_{n+1}(S,\theta)\}_{n=\underline{\eta}}^{\overline{\eta}}$ and retirement age $\hat{n}(S,\theta)$ solve an agent's problem in (7) – (13).

2. For each ability group θ, the fraction of agents making schooling choice S is equal to $p(S|\theta)$, as determined by agents' decision rule on schooling in (14).

3. Prices $\{r, R^1, R^2\}$ are consistent with aggregate quantities $\{\bar{K}, \bar{H}^1, \bar{H}^2\}$ of inputs in goods production under perfect competition, i.e., prices satisfy (15) – (18).

4. The capital market clears,

$$\kappa \bar{K} = \sum_{\theta} \sum_{S} \frac{p(S|\theta)}{4} \sum_{n=\underline{\eta}}^{\overline{\eta}} \hat{K}_{n+1}(S,\theta). \tag{19}$$

5. Both labor markets, $S = 1, 2$, clear,

$$\bar{H}^S = \sum_{\theta} \frac{p(S|\theta)}{4} \sum_{n=\underline{\eta}+4(S-1)}^{\hat{n}(S,\theta)-1} e(n)\left(1 - \hat{I}_n(S,\theta)\right)\hat{H}_n(S,\theta). \tag{20}$$

6. Goods market clears,

$$\sum_{\theta}\sum_{S} \frac{p(S|\theta)}{4} \sum_{n=\underline{\eta}}^{\overline{\eta}} \hat{C}_n(S,\theta) \;+\; \sum_{\theta} \frac{p(2|\theta)}{4} \cdot 4\zeta \;+\; G = \tilde{F}(\bar{H}^1, \bar{H}^2, \bar{K}). \tag{21}$$

7. Government policy $\{\tau_l, \tau_k, \tau_p, \eta_p, P, G\}$ satisfies the government budget constraint,

$$\begin{aligned} G + (1-\tau_l)\sum_{\theta}\sum_{S} \frac{p(S|\theta)}{4}\Big(\overline{\eta} - \max\{\eta_p, \hat{n}(S,\theta)\} + 1\Big)P \\ = (\tau_l + \tau_p)(R^1\bar{H}^1 + R^2\bar{H}^2) + \tau_k r\bar{K}. \end{aligned} \tag{22}$$

All parameterizations considered in this paper satisfy regularity requirement (i). As explained in Section 3.1, under our parameterizations in which an equilibrium after-tax interest rate is greater than the subjective rate of discounting, in combination with the other four features mentioned above that prevail in our model, front-loaded labor supplies are optimal. But regularity condition (ii) is not satisfied in some of our tax experiments below when agents with the same abilities and schoolings become indifferent between career strategies that differ in terms of retirement ages and human capitals accumulated on the job, as described in Section 5.2. Our definition of a stationary equilibrium can be extended to allow for such indifference by introducing equilibrium fractions of otherwise identical agents who nevertheless choose differing career strategies that yield the same lifetime utilities. Equilibrium fractions of agents who choose different strategies are pinned down in a stationary equilibrium.

3.4 Calibration

To set the stage for our policy experiments, we calibrate our time-averaging model to mimic outcomes of HLT's analyses. Appendices 2.1 and 2.2 describe how we solve the model via a shooting algorithm and how we compute prices. To retain most of HLT's parameterization while adding essential time-averaging features, we deploy four steps.

As a first step, except for us above having set both γ and ρ_2 to zero, we assume that HLT's remaining parameterization in Table 1 also applies to our time-averaging model. Furthermore, we assume that our equilibrium interest rate and skill prices will equal HLT's in Table 3. Subject to those restrictions, we explore alternative configurations of core features of our time-averaging model – disutility of work at the extensive margin, depreciation of human capital in old age, and a social security program – that induce all agents to retire at the official retirement age 65.

We posit a payroll tax rate $\tau_p = 0.10$ and target a social security benefit $P = 8$ that is equivalent to 40% of average earnings. This constant benefit implies a progressive replacement rate.[12] Parameters ϕ_1 and ϕ_2 in (6) that determine age-dependent efficiency units of human capital, in conjunction with the disutility B of working, can then reside in a nontrivial subspace that induces all agents to retire at age 65. For example, consider a degenerate case that lets ϕ_1 go to infinity and sets $\phi_2 = 65$ so that human capital produces efficiency units one-for-one until age 65, and then drops to zero. For sufficiently low values of B, all agents choose to retire at age 65, as in HLT's analyses. But because we want agents to be capable of working beyond age 65, we set the "slope" coefficient $\phi_1 = 0.2$ and move the inflection point to $\phi_2 = 75$ in (6). This generates a smoother decline in the efficiency units of human capital and the depreciation becomes noticeable only at ages in the 60s, as depicted by the solid line in Figure 1. Finally, given this suite of parameter values, we set the disutility of working $B = 0.8$ that lies in a mid-range of values for which all agents choose to retire at age 65. For a sensitivity analysis of these baseline parameters, see Section 8.1.

In our second step, we calibrate probability distributions for nonpecuniary costs ϵ of attending college. We have altered HLT's specification by denominating that shock in utils instead of dollars. Consequently, we want to calibrate ability-specific means μ_θ and a common standard deviation σ to be compatible with these units. Ideally, we would want to use HLT's college enrollment rates as targets for our calibration, but HLT do not report those rates. So instead we use the same source that gave us a value ζ of annual college tuition. For a given value of σ, we calibrate ability-specific means μ_θ to target college enrollment rates of different ability groups in Taber (2002, Table 1): we report these outcomes in our model's

[12]Relative to labor earnings at the time of retirement, the social security benefit $P = 8$ corresponds to a replacement rate of 52.2% (27.8%) for the lowest (highest) earner in our model, which reflects that the lowest earner earns about half of as much as the highest. In comparison, according to the OECD study of Queisser and Whitehouse (2006), the average gross replacement rate in the U.S. social security system for an individual with the average earnings is 38.6%, while it is 49.6% and 28.1% for individuals with half and twice the average earnings, respectively.

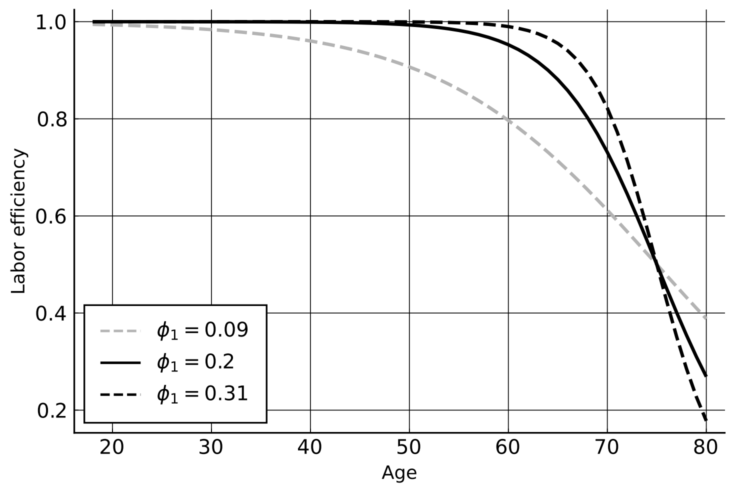


Figure 1: Age profile of the conversion of one unit of human capital into efficiency units.

baseline steady state in Table 3. While postponing calibration of σ until our last step, our calibration of μ_θ lets us recover HLT's aggregate composition of high school and college graduates in the labor force, outcomes that will support skill prices close to those of HLT under our maintained assumption of an equilibrium interest rate equal to HLT's.

Our third step ensures that our equilibrium interest rate agrees with HLT's. We back out a physical capital stock that supports that outcome and compute agents' net savings. The ratio of the latter to the former is our calibrated value of a new parameter κ, the fraction of physical capital held by agents who live inside our model; remaining capital is held by 'investors' who live outside our model. Since we are matching HLT's outcomes with a model that includes a social security program, we should anticipate that our agents' private net savings fall far short of a physical capital stock required to generate the HLT's equilibrium interest rate. Not surprisingly, our calibrated value is $\kappa = 0.388$.

Our fourth step calibrates σ by using both steady states computed by HLT – a baseline steady state to describe the early 1970s and a future steady state to which their economy converges after a 30-year period of skill-biased technological change, as described in Section 2.1. We calibrate σ by targeting the difference in the relative skill price ratio between the two steady states. HLT's permanent shift in technology toward skilled labor increases the price of college human capital relative to the price of high school human capital by 8%. Our calibrated value of σ comes from our finding of a positive relationship between that parameter and the implied increase in the relative skill price. A larger dispersion in the idiosyncratic nonpecuniary cost of attending college means that college attendance rates respond less to an increase in the relative skill price, so it takes a bigger change in skill prices to elicit reallocation from high school human capital to college human capital in response to skill-biased technological change.

Tables 1, 2 and 4 list a complete parameterization of our time-averaging model. Besides the calibration issues discussed above, we make minor adjustments to HLT's production parameters $\{a_1, a_2, a_3\}$ in Table 1 in order exactly to target prices in HLT's baseline steady state. These minor adjustments are quantitatively insignificant in the sense that if we instead had adopted HLT's parameter values for $\{a_1, a_2, a_3\}$ in Table 1, the effects on our reproduction of HLT's outcomes in the last two columns of Table 3 would be unnoticeable, as would outcomes from the policy analyses below.

3.5 Our model mimics outcomes in HLT's

The last two columns of Table 3 show that our time-averaging model does a good job of approximating HLT's baseline steady state and also HLT's second steady state after skill-biased technological change. To help understand the sources of these close approximations, we describe how primitives inherited from the HLT model interact with key new features of endogenous career length and a social security system that we put into our time-averaging model.

In contrast to HLT's assumption of inelastic labor supply until exogenous retirement at age 65, career length and time of retirement are endogenous in our model. As in HLT, after agents draw their abilities and their nonpecuniary costs of attending college, they face no risks, so a competitive market in risk-free one-period loans is enough to ensure efficient on-the-job human capital investments *conditional* on a choice of career length. For a given career length, optimal on-the-job human capital investments are those that maximize the present value of labor earnings. Hence, those investment decisions are decoupled from a worker's choice of when to consume goods. But the choice of career length is *not* decoupled in that way because of how it determines utility derived from leisure in retirement. Furthermore, since the after-tax market interest rate is higher than the subjective discount rate in both HLT's analysis and ours, an agent prefers to front-load lifetime labor supply in order to earn rate of return on savings accumulated from past wages that exceeds his time preference discount rate. All of these considerations affecting endogenous labor supplies become moot in our baseline steady state because the social security system induces agents to put themselves on a corner that involves working until the official retirement age 65. As a consequence, all workers in our baseline steady state find it optimal to maximize their lifetime after-tax labor incomes before retiring at age 65, just as in HLT's model. Since we adopt the same human capital technologies as HLT, it is not surprising that human capital investments and lifetime earnings profiles in our baseline steady state closely approximate HLT's.

There is no presumption that our time-averaging model can also do a good job of approximating HLT's second post-skill-biased technological change steady state. Thus, we might anticipate that in our model some agents might leave the corner solution that tells them to retire at the official retirement age 65. Other differences might arise because other auxiliary assumptions that influence the aggregate stock of physical capital. Thus, recall HLT's constant per capita lump-sum transfer X from the old to the young that they adjust to target their baseline steady-state capital-output ratio; we instead hit that target by assuming that agents who live inside our model hold only a fraction κ of the equilibrium capital stock and are entitled to social security benefits at age 65 if they have chosen to retire then. Despite these important difference in primitives, we find that our time-averaging model does a good job of approximating HLT's post-technical-change steady state. A critical outcome in our model is that agents continue to retire at the official retirement age 65. So once again, with the same human capital technologies and now also having experienced the same technological change as the agents in HLT's model, agents in our model enroll in college and invest in human capital at the same rates as do the agents in HLT's model.

Thus, HLT could have used our time-averaging model with endogenous retirement to get their same quantitative findings. Our model with a social security system thus offers a structural rationalization for retirement outcomes that HLT hard-wired.[13] Furthermore, since our agents choose a corner solution that tells them to retire at the official retirement age, and if we were explicitly to model labor indivisibilities that induce primary workers to choose to work full time rather than part time, as we shall discuss in Section 8.2, HLT's inelastic labor supply emerges in our model automatically.

Although these structural differences between our model and HLT's make no practical differences for the issues studied by HLT, they do matter when we study social security reforms and possible employment effects of increasing the labor income tax studied by Prescott (2002).

[13] A common practice is to justify exogenously requiring agents to retire at age 65 by alluding informally to rules of a social security system that used to prevail before recent reforms designed not to discourage working beyond age 65.

4 Social security reform

As our first policy experiment we assume that all workers receive social security benefits from age 65 regardless of when they choose to retire. Under our assumption that agents face no risks after beginning-of-life abilities and nonpecuniary costs of attending college are realized and our assumption that agents can borrow and lend at a risk-free interest rate, the only consequence of this reform is to remove an implicit tax on working after the official retirement age.

The first two columns in Table 5 show retirement ages under this reform when prices are kept fixed and in a general equilibrium, respectively, where age i (j) in entry 'i/j' is the retirement age of a high school (college) worker. In the first column, both the interest rate and skill prices are kept constant at their values in the baseline economy; in addition, we also froze the ability composition when we computed an average retirement age. Since at fixed prices before the reform they had chosen a corner solution at the official retirement age 65, all workers choose to extend their career lengths. While high school workers increase their career lengths on average by 2.4 years, college educated workers increase their average retirement age of by 7.6 years to 72.6. These changes are attenuated in a general equilibrium: high school workers actually choose to retire *early*, on average one year before the official retirement age. This experiment thus unleashes countervailing forces that we shall unbundle by successively perturbing distinct aspects of the environment.

There are basically three reasons that high school workers retire earlier than college workers. First, the social security system redistributes from high-ability to low-ability agents. Since all workers pay the same proportional payroll tax on labor income, the equal social security benefit awarded to all retirees means that agents with low earnings receive more than they pay into the system. Income effects of that net transfer to low-ability agents, who are more likely to be high school workers, induces them to supply less labor. Our first perturbation is designed to quantify this effect: we simply remove the social security system so that agents must save in order to consume when they aren't working during their retirements. Except for recalibrating the fraction of the economy's capital stock that is held by agents who live inside our model,[14] we retain our parameterization of the baseline

[14]As discussed in Section 3.4, the pay-as-you-go social security system in the baseline economy significantly suppresses agents' private retirement savings. Hence, the removal of social security in our perturbed model increases those savings. For reasons of comparability, we therefore recalibrate the fraction of the capital stock held by the agents in the model to preserve HLT's calibration target of a capital-output ratio equal to 4. Doing that calls for an increase in the fraction of the capital stock held by the agents in the model from 0.388 to 0.767. We note that with that recalibration we can replicate both the baseline and the second

Table 5: Retirement ages [high school/college workers] with and without social security

Ability	**Social security reform**		**No social security**	
	fixed prices⋆	general eq.	new baseline†	perturbed techn.‡
$\theta = 1$	66.7 / 72.8	63.3 / 70.9	64.0 / 71.0	66.1 / 70.3
$\theta = 2$	67.7 / 72.8	64.1 / 70.9	64.6 / 71.0	66.0 / 70.3
$\theta = 3$	68.2 / 72.5	64.6 / 70.6	65.1 / 70.7	65.5 / 70.1
$\theta = 4$	68.2 / 72.7	64.5 / 70.8	65.0 / 70.8	65.8 / 70.2
Average	67.4 / 72.6	63.9 / 70.8	64.5 / 70.8	65.9 / 70.2

⋆ Both the interest rate and skill prices are kept constant at the values of the baseline economy, as well as the ability composition when computing an average retirement age.

† In the new baseline of an economy without social security, the fraction of the capital stock held by the agents in the model are recalibrated so as to maintain a capital-output ratio equal to 4.

‡ Human capital technologies are perturbed so that high school and college workers of the same ability have the same technology and initial human capital, equal to those of the latter workers.

economy and compute a steady state without social security. The third column of Table 5 shows that workers who had chosen to retire the soonest under the social security reform – high school workers of the lowest ability – increase their career lengths marginally more than other workers in response to removing all social security benefits.

Second, recall Ljungqvist and Sargent's (2014) finding that the more elastic is an earnings profile to accumulated time worked, the longer is a worker's optimal career. We anticipate that the same force is present in our model here with its Ben-Porath human capital technology. To quantify this force, we use our next perturbation of the economy without social security, which is to endow high school workers with the same human capital accumulation technology that college workers have. This technology has higher returns on time devoted to human capital accumulation on the job than does the one that high school workers ordinarily have in our model. Thus our perturbed human capital technology parameters, denoted $\{\hat{\alpha}_S, \hat{\beta}_S, \hat{A}^S(\theta), \hat{H}^S(\theta)\}$, satisfy $\hat{\alpha}^1 = \hat{\alpha}^2 = \alpha^2$, $\hat{\beta}^1 = \hat{\beta}^2 = \beta^2$; and, for each θ,

steady state of HLT in our perturbed model, conditional on adding the restriction that agents cannot work beyond age 65.

$\hat{A}^1(\theta) = \hat{A}^2(\theta) = A^2(\theta)$, and $\hat{H}^1(\theta) = \hat{H}^2(\theta) = H^2(\theta)$. We keep all other parameters the same as in the preceding perturbed economy without social security. As the fourth column of Table 5 shows, letting high school workers have college workers' human capital technology induces them to choose longer careers in the new steady state.

Third, we ascribe to yet another effect of time averaging an approximate 4-year difference in career lengths between high school and college workers in the fourth column of Table 5. If to the bare-bones time-averaging framework of Ljungqvist and Sargent (2006, 2014), we were to add the assumption that an initial apprenticeship period Z of a labor market career yields no labor earnings but is simply required before starting gainful employment, at an interior solution an optimal career length is the sum of Z and a corresponding optimal career length in an economy without that apprenticeship.[15] Evidently, this force is at work in our present time-averaging version of an HLT framework. Thus, in the perturbed model without social security and identical human capital accumulation technologies for high school and college workers of the same abilities, if an agent finds it optimal to acquire a college degree, he apparently treats the years in college as a fixed requirement and simply tack them on to the length of time spent working that he would have chosen if he had instead gone to work straight out of high school. Thus, regardless of the schooling choice, the optimal time spent actually working depends only on the fixed disutility of working and the human capital accumulation technology that influences how agents choose to shape their earnings profiles. Furthermore, because workers' preferences are compatible with balanced growth, absolute wage levels *per se* do not affect labor supply decisions: income and substitution effects cancel.

[15] Consider the simplest version of the time averaging model in continuous time studied by Ljungqvist and Sargent (2006, 2014). Over a deterministic lifespan of unit length, an agent maximizes lifetime utility $\int_0^1 e^{-\varrho t}[\log(c_t) - Bn_t]\,dt$ subject to a present value budget constraint $\int_0^1 e^{-rt}[wn_t - c_t]\,dt \geq 0$. Under the assumptions that the agent's subjective discount rate ϱ and the market interest rate r are the same and equal to 0, the optimization problem reduces to

$$\max_T \,[\log(c) - BT] \quad \text{subject to } \; c = wT, \;\; c \geq 0, \;\; T \in [0,1]\,,$$

i.e., the agent chooses a constant consumption stream c and a fraction T of the lifespan devoted to working. At an interior solution, the optimal career length is $T = 1/B$.

Perturb the model by assuming that an initial interval $Z \in [0,1)$ of a career yields no labor income so that Z is time spent to acquire necessary training. The optimization problem then becomes

$$\max_T \,[\log(c) - BT] \quad \text{subject to } \; c = w \cdot \max\{T - Z, 0\}, \;\; c \geq 0, \;\; T \in [0,1]\,.$$

At an interior solution, the optimal career length is $T = 1/B + Z$.

5 Taxation and spending

To study how responses of labor supplies to increases in labor income tax rates depend on how the government allocates tax revenues, we now conduct a tax experiment along the lines of Prescott (2002).[16] Prescott argued that most tax revenues are spent for goods and services that substitute perfectly for private consumption. Therefore, he modelled government expenditures as lump-sum transfers to households. Temporarily embracing Prescott's assumption, we explore the effects of increasing our labor income tax rate from its baseline parameterization of $\tau_l = 0.15$. At higher tax rates $\tau_l > 0.15$, a fraction $(\tau_l - 0.15)/\tau_l$ of all tax revenues raised by levying the tax rate τ_l on labor income is returned to each agent as a lump-sum transfer. Two remarks are pertinent. First, we note that the lump-sum transfer includes no tax revenues from levying the tax rate τ_l on social security benefits. Second, because the baseline tax rate is 15 percent, revenues from the higher tax rate τ_l are in general higher than the change in total revenues from labor income taxation. Smaller tax revenues can be expected to be raised from the 15-percent baseline portion of the labor income tax whenever a higher tax rate τ_l causes labor supplies and consequently aggregate labor income to shrink. Such losses of baseline tax revenues do not affect our calculation of a lump-sum transfer. Our tax experiment risks bankrupting the government, so as we raise the labor income tax rate, we must verify that the sum of labor income tax revenues not handed back as lump-sum transfers and the revenues from the capital tax and the payroll tax are sufficient for the government to finance social security benefits.

5.1 Aggregate outcomes

The solid line in Figure 2(a) depicts a Laffer curve with respect to tax revenues that are actually handed back as lump-sum transfers, expressed in per capita terms, i.e., lump-sum transfer at different tax rates $\tau_l \geq 0.15$. The Laffer curve peaks at labor tax rate $\tau_l = 0.54$. By way of contrast, the dashed line in Figure 2(a) depicts the Laffer curve that would prevail if tax revenues were instead to be used to finance government expenditures that are *not* good substitutes for private consumption. The two Laffer curves ratify Prescott's (2002, p. 7) assertion that "the assumption that the tax revenues are given back to households either as transfers or as goods and services [that are good substitutes to private expenditures] matters. If these revenues are used for some public good or are squandered, private consumption will fall, and the tax wedge will have little consequence for labor supply." Nevertheless, because

[16] Appendix B.3 describes computational details about Laffer curves.

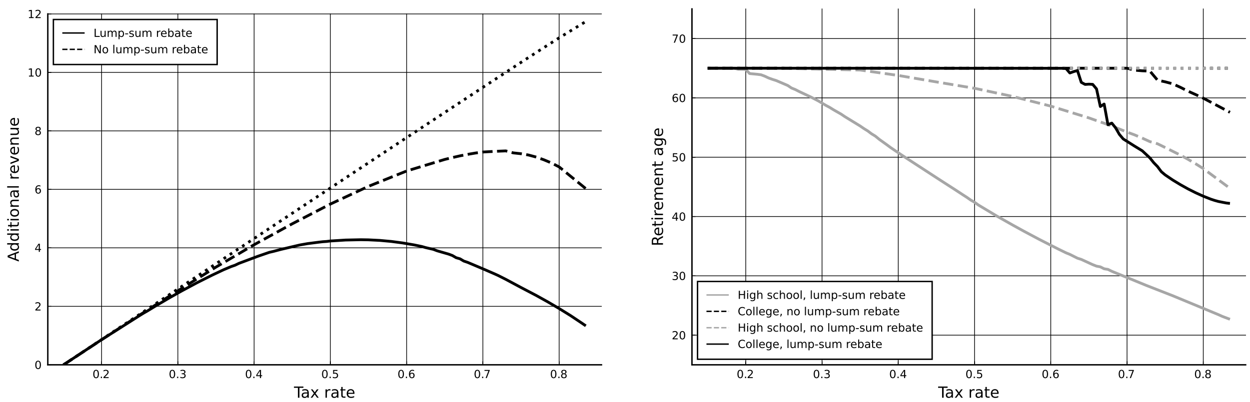


Figure 2: Laffer curves and retirement ages as a function of the labor tax rate, and whether or not tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as equal lump-sum transfers to agents. In Panel A, the solid and the dashed line is the Laffer curve with and without lump-sum rebates, respectively. In Panel B, the dark and light solid (dashed) line is the average retirement age of college and high school workers, respectively, with (without) lump-sum rebates. In both panels, the dotted line depicts partial-equilibrium outcomes in the economy without lump-sum rebates, when the interest rate is kept constant at the baseline equilibrium rate. Notice that the single dotted line in Panel B says that all workers retire at age 65 throughout the computed tax range.

capital formation is affected in a general equilibrium, the tax wedge on labor income brings distortions that increase along with the tax rate. Thus, the dashed Laffer curve peaks at tax rate $\tau_l = 0.73$ for the economy without lump-sum rebates. To examine how much of the distortions operate through capital formation, the dotted line in Figure 2(a) is the Laffer curve for a small-open version without lump-sum rebates when the interest rate is held constant at the baseline equilibrium rate.

Tax distortions on labor supply manifest themselves as changes in career lengths, college enrollment rates, and how workers allocate their time between working and accumulating human capital on the job. Figure 2(b) depicts effects of the labor income tax rate on retirement ages. Dark and light solid lines show average retirement ages of college and high school workers, respectively, in the economy with lump-sum rebates. At the baseline tax rate $\tau_l = 0.15$, outcomes are those of the baseline steady state of Section 3.5. High school workers are the first to start retiring early in response to higher tax rates and associated lump-sum transfers, while college workers retire early only when tax rates are higher than $\tau_l = 0.60$. Counterparts in the form of dashed lines show that these effects on retirement ages are attenuated in an economy without lump-sum rebates. Furthermore, in the small-open version of our no-lump-sum-rebate economy, over the computed tax range all workers

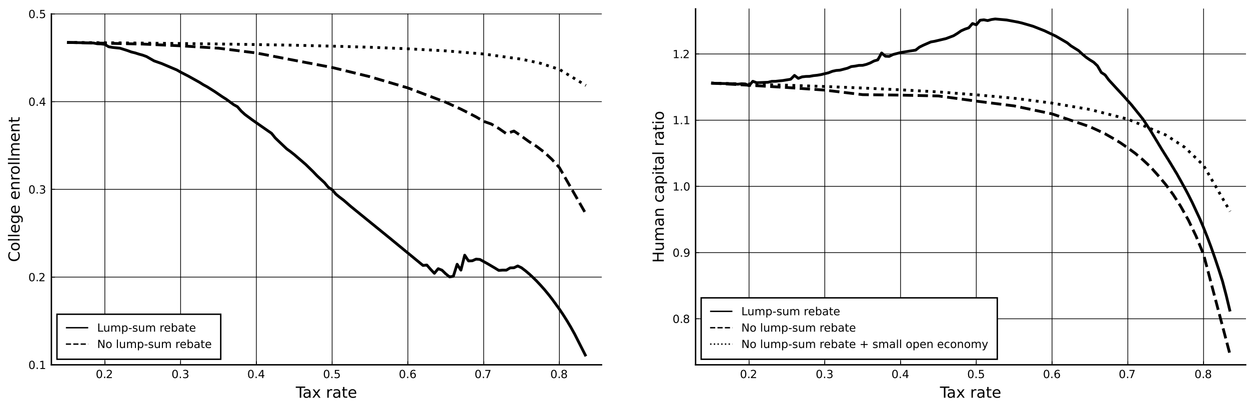


Figure 3: College enrollment (Panel A) and ratio of college to high school human capital in the production of goods (Panel B) as a function of the labor tax rate, and whether or not tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as equal lump-sum transfers to agents. In both panels, the solid and the dashed line displays outcomes with and without lump-sum rebates, respectively. The dotted line depicts partial-equilibrium outcomes in the economy without lump-sum rebates, when the interest rate is kept constant at the baseline equilibrium rate.

continue to work until age 65, i.e., the dotted line in Figure 2(b).

In the economy with lump-sum rebates, shortenings of high school workers' career lengths and the labor supplies of college workers that are less sensitive to tax rate increases are reconciled in a general equilibrium through a falling college enrollment rate, as indicated by the solid line in Figure 3(a). The depressed labor supply of the average high school worker brings adjustments of equilibrium prices that induce more agents to commence working with high school degrees rather than going to college. Workers' choices of how much human capital to accumulate on their jobs also influence aggregate quantities of college and high school human capitals employed to produce goods. Connections between these choices and workers' decisions about when to retire result in the almost flat segment in college enrollment at the upper range of taxes, to be discussed in Section 5.2. Taken together, as shown by the solid line in Figure 3(b), these forces on labor supplies result in a ratio of aggregate college to high school human capital devoted to producing goods that initially increases with the tax rate, peaks around the same tax rate as does the Laffer curve in Figure 2(a), and then starts falling ever more rapidly, eventually in tandem with a sharp decline in the college enrollment rate in Figure 3(a).

The dashed line in Figure 3(a) shows that effects on college enrollment are attenuated in the economy without lump-sum rebates and, according to the dotted line, are even more attenuated in a small-open economy version. In the absence of lump-sum rebates, the ratio of college to high school human capitals used to produce goods decreases monotonically as the tax rate rises. To gain a better understanding of these outcomes, we inspect the equilibrium interest rate and prices of the two types of human capital.

In the small-open version of the economy without lump-sum rebates, the interest rate is kept constant at the baseline equilibrium rate 0.059 ($= 0.05/(1 - \tau_k)$), as indicated by the dotted line in Figure 4(a). According to the dashed line in Figure 4(b), increases in the labor tax rate are accompanied by a capital inflow from abroad, expressed as a fraction of the equilibrium capital stock at each tax rate. Here is what happens. Consider a standard laissez-faire growth model with preferences consistent with balanced growth, as is true for our utility function (4). A permanent decline in a multiplicative productivity parameter would cause a proportional reduction in the wage rate that would leave steady-state labor supply unchanged as substitution and income effects cancel. Other things equal, a similar cancellation of income and substitution effects would occur if the take-home wage rate were instead to be reduced by levying a proportional labor income tax (with all tax revenues being squandered) and hence, steady-state labor supply would remain unchanged. But capital formation rates are not equal across the two settings. While a multiplicative deterioration of the production function would leave all relevant equilibrium ratios unchanged including the fraction of agents' income devoted to investments that sustain the new steady-state capital stock, such invariance of equilibrium ratios would not prevail if the reduction of agents' take-home wage rate came about because of a proportional labor tax. In particular, since the production technology has not changed, the capital stock would need to stay unchanged in order to justify our temporary assumption of an unchanged before-tax wage rate upon which the cancellation of substitution and income effects hinges. But the capital stock would have to change because the investments required to sustain that unchanged capital stock constitute a larger share of agents' now depressed after-tax income. However, if we were to assume a small-open economy with an interest rate held constant at the economy's steady-state rate prior to the imposition of the proportional labor tax, outcomes would indeed be the same as if there had been a multiplicative deterioration of the production function. In the words of Prescott (2002, p. 7) above, "private consumption will fall, and the tax wedge will have little consequence for labor supply." Indeed, there would be no effect at all on labor supply because the capital flowing into the economy from abroad would completely make up for the shortfall in domestic savings at the unchanged interest rate. This reasoning explains the capital inflow along the dashed line in Figure 4(b) that sustains the constant interest

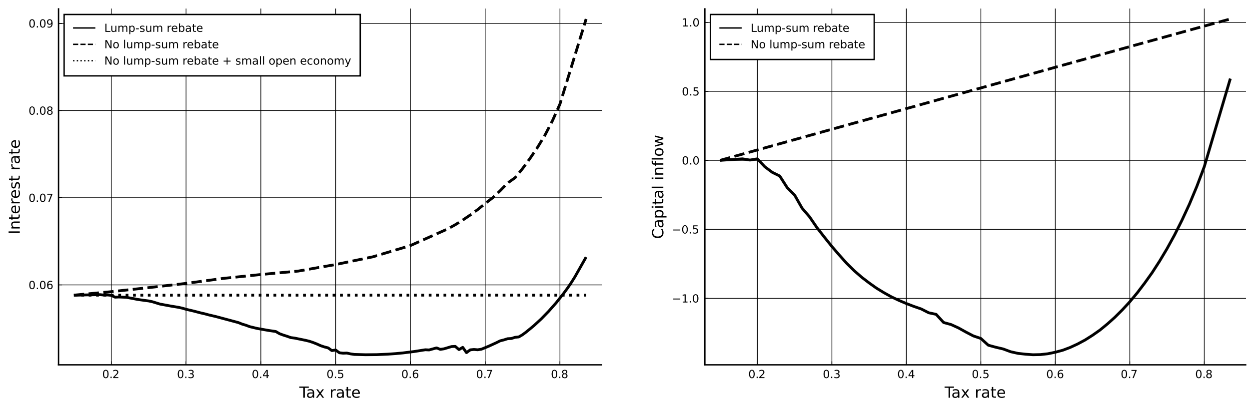


Figure 4: Interest rate (Panel A) and capital inflow into small open economy (Panel B) as a function of the labor tax rate, and whether or not tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as equal lump-sum transfers to agents. In both panels, the solid and the dashed line displays outcomes with and without lump-sum rebates, respectively. The dotted line in Panel A depicts the constant interest rate in a small open economy, i.e., our baseline equilibrium interest rate.

rate indicated by the dotted line in Figure 4(a), which in turn rationalizes the nearly linear dotted Laffer curve in Figure 2(a) that implies an approximately unchanged labor supply over the depicted tax range.[17]

By applying this reasoning to our economy without lump-sum rebates, the monotonically increasing interest rate along the dashed line in Figure 4(a) follows immediately. Thus, to offset the capital inflow that occurs in the small-open economy version, the general-equilibrium interest rate has to rise in order to increase agents' savings and to reduce firms' demand for capital services. While our assumption of preferences consistent with balanced growth props up labor supplies in response to a proportional labor tax when tax revenues are squandered or spent on public goods that are not good substitutes for private expenditures, agents' propensities to save enough to maintain the capital stock are now suppressed by their diminished after-tax incomes. So the interest rate must rise.

The picture gets more complicated if tax revenues are returned lump-sum. Lump-sum rebates suppress the income effect of the proportional labor tax and give substitution effect free rein to reduce labor supply. As for the equilibrium interest rate, we make two observations. First, in the lump-sum-rebate economy, countervailing forces no longer operate in

[17]In contrast to our account of outcomes in a standard growth model, only an approximate invariance of labor supply to a proportional labor income tax prevails in the small-open economy version of our model. This comes from our added feature of a schooling choice with a nonpecuniary cost of attending college, as will be explained below.

the capital market. Instead, since the labor tax rate now has a strong suppressive effect on labor supply, the lower savings of workers with their smaller incomes and the lower demand for capital services by firms producing less ouput go hand in hand as the tax rate increases. Consequently, we can no longer argue that the interest rate must unequivocally rise in response to a higher tax rate as it does in the no-lump-sum-rebate economy. Second, lifecycle savings forces lead us to anticipate that the interest rate might now fall in response to a higher tax rate. This is because, while shortened career lengths reduce agents' labor income, it also increases their motive to save parts of that labor income early in life to prepare for more years of not working when retired. This exerts a downward pressure on the interest rate. However, this force should dissipate at high enough tax rates when the lump-sum transfer becomes large relative to net-of-tax lifetime labor earnings. The transfer effectively shifts agents' disposable incomes forward over their lifetimes, supplementing their net-of-tax social security benefits in old age. The resulting decline in agents' demand for private savings to finance future consumption exerts an upward pressure on the interest rate in response to any further increases in the tax rate. These forces play out in Figure 4(a) where the solid line shows that the interest rate is a U-shaped function of the tax rate. The corresponding outcome for capital flows in the small-open version of the lump-sum-rebate economy is depicted by the solid line in Figure 4(b) .

Figure 5 depicts the relative price of college and high school human capitals – the skill premium; it qualitatively resembles Figure 4(a) that shows the interest rate. In the economy without lump-sum rebates (dashed lines), that both the skill premium and the interest rate increase in the labor tax rate emerges from countervailing forces on the attractiveness of attending college. While a higher skill premium increases the returns to a college degree, a higher interest rate lowers it because costs in the form of tuition and four years of lost labor earnings are incurred upfront, while future higher earnings as a college graduate are subject to a higher discount rate. These countervailing forces come from how the interest rate affects investments in both physical and human capital and how general equilibrium reconciles outcomes with price adjustments. Our focus above was on how movements in the interest rate can offset what otherwise would be capital flows in a small-open version of the economy. As a force that moderates impacts of the interest rate on human capital investments, movements in the skill premium create incentives to attend college. To close the general-equilibrium circle, prices of human capital are equated to marginal products of college and high school human capitals, so the relative price in Figure 5 should indeed be an inverse image of the ratio of those two inputs in production that is depicted in Figure 3(b).

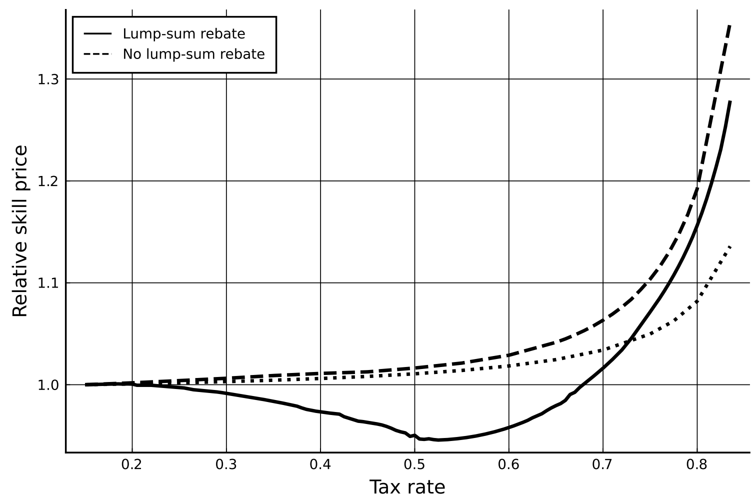


Figure 5: Relative price of college to high school human capital as a function of the labor tax rate, and whether or not tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as equal lump-sum transfers to agents. The solid and the dashed line displays outcomes with and without lump-sum rebates, respectively. The dotted line depicts partial-equilibrium outcomes in the economy without lump-sum rebates, when the interest rate is kept constant at the baseline equilibrium rate.

Reverse forces are evidently present in the economy with lump-sum rebates (solid lines) as both the interest rate and the skill premium at first decrease with increases in the labor tax rate in Figures 4(a) and 5, respectively. A falling interest offsets what otherwise would provoke a capital outflow in the small-open economy version. Evidently, this falling interest rate stimulates investments in college human capital so much that the skill premium falls. For high enough tax rates, both the skill premium and the interest rate will increase in the economy with lump-sum rebates, as they do in the economy without lump-sum rebates. While the same prime force increases the interest rate in these two economies, paths of causation differ. The prime force is that agents in both economies are deprived of resources to invest in physical capital. This is obvious for the economy without lump-sum rebates since agents are immediately left with lower after-tax incomes. In the economy with lump-sum rebates agents also eventually end up with reduced resources to invest in physical capital, but now this happens because the tax-transfer policy is so powerful in disincentivizing labor supplies that the economy's output plummets, and the transfers reshuffle agents' disposable incomes over time so that little is available to be invested in physical capital, as described above. These different paths of casuation manifest themselves in the Laffer curves in Figure

2(a). The solid Laffer curve for the economy with lump-sum rebates approaches zero, reflecting vanishing aggregate production. In contrast, the dashed Laffer curve for the economy without lump-sum rebates is higher up with plenty of tax revenues for the government to squander or to finance public goods that are not close substitutes to private expenditures, while agents are deprived of resources for private consumption and investments in the capital stock.

By construction, the interest rate does not change in the small-open version of the economy without lump-sum rebates, but we notice an eventual increase in the relative price of college to high school human capital (dotted line) in Figure 5. The skill premium compensates a marginal agent for investing in a college degree. Besides payments of tuition, as the idiosyncratic nonpecuniary cost ϵ of attending college becomes larger relative to the diminished lifetime utility of consumption when agents are deprived of resources, the skill premium must increase to compensate agents who choose to become college graduates rather than to seek employment as high school workers.

We turn next to the decisions by individual agents that underlie these aggregate outcomes.

5.2 Individual outcomes

In this section, we focus on agents' behavior and outcomes in the economy with lump-sum rebates: tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are returned as lump-sum transfers to all agents.

As in HLT's model, heterogeneity in our baseline equilibrium occurs in the form of 8 distinct classes indexed by 8 pairs consisting of an agent's endowed ability, which belongs to one of four possible ability groups, and an agent's schooling level, which can be either a high school or a college graduate. At high enough tax rates, it is possible that some agents with the same ability and schooling are indifferent between career strategies that differ in terms of retirement age and human capital accumulated on the job. We describe this last heterogeneity below.

The data indicate a high correlation between levels of ability as measured by test scores and an individual's propensity to acquire a college degree. Different rates of college enrollment by ability in our baseline steady state, as reported in the third column of Table 3, appear in Figure 6 at baseline labor tax rate $\tau_l = 0.15$. How those college enrollment rates respond to changes in the tax rate depend on our having embraced HLT's estimates of ability-specific distributions of nonpecuniary costs of attending college and ability-specific

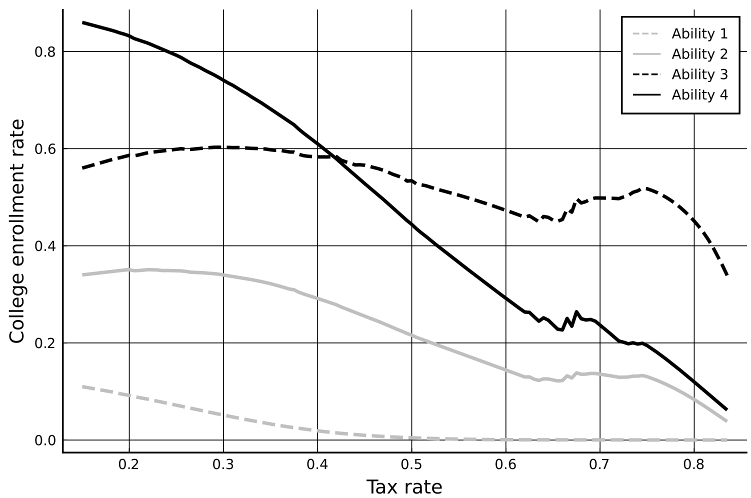


Figure 6: College enrollment rate by ability group as a function of the labor tax rate. Tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as equal lump-sum transfers to agents.

initial endowments of high school and college human capital. To match evidence of high earnings and steep earnings profiles of college workers of the highest ability group 4, HLT inferred that their initial endowment of college human capital is high and that they also reap a high return from additional investments in human capital on the job. To explain how a non-negligible fraction of agents in ability group 4 nevertheless chooses to supply labor as high school workers who actually earn less than high school workers from the next highest ability group 3, HLT imputed to them a high average disutility of attending college. Thus, despite large relative and absolute advantages of becoming college workers, their high disutilities of attending college cause 14 percent of agents in ability group 4 to become high school workers in the baseline steady state. Those high disutilities of attending college explain why the college enrollment rate of ability group 4 falls most sharply in response to higher labor income taxation in Figure 6. At the opposite end of the ability spectrum, only ability group 1 has a higher average disutility of attending college than ability group 4. Ability group 1 also has by far the lowest endowment of college human capital. Consequently, only 11 percent of agents in ability group 1 attend college in the baseline steady state; that fraction falls to less than one percent at labor tax rates above 0.45 in Figure 6.

In addition to depending on college enrollment rates, supplies of high school and college human capital also depend on how much human capital agents accumulate on the job. Figure 7 shows end-of-life human capitals of high school workers of the four ability groups

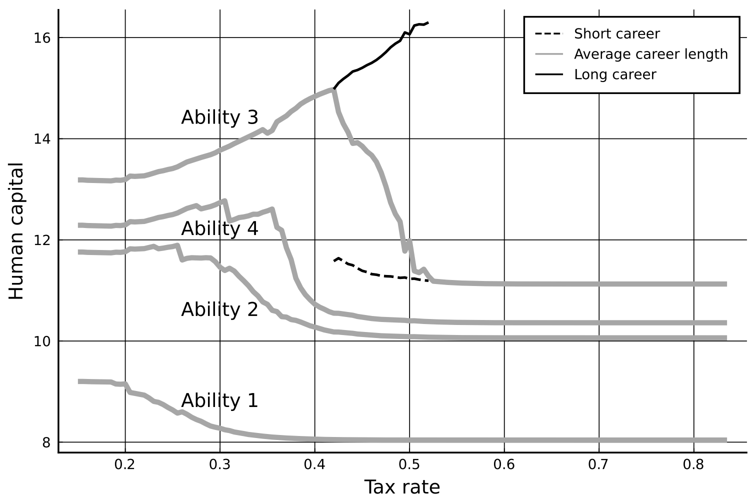


Figure 7: End-of-life human capital of high school workers by ability group as a function of the labor tax rate. Regarding ability group 3, over the tax range 0.42–0.52, workers are indifferent between accumulating relatively high and low human capital, as depicted by the dark solid and dark dashed line, respectively, whereas the light solid line inbetween shows a worker's average human capital in an equilibrium. Tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as equal lump-sum transfers to agents.

as functions of the labor tax rate. These are measured in raw units of human capital, not age-dependent efficiency units; given the absence of depreciation from (2), an agent's end-of-life human capital is the maximal human capital attained over his labor market career. For the lower ability groups 1 and 2, human capital accumulation tends to decrease as the tax rate increases – an adverse effect of taxation that could have been anticipated. Although that outcome eventually also prevails for ability groups 3 and 4, initially we see opposite outcomes whenever end-of-life human capital increases with increases in the tax rate in Figure 7. This pattern reflects general equilibrium effects when the falling supply of high school human capital of workers from the lower ability groups provokes higher supplies from the two higher ability groups. Initially, along the solid line in Figure 4(a), a falling interest rate motivates those higher ability workers to accumulate more human capital. In contrast, a high school worker from the lowest ability group 1 monotonically reduces human capital in response to a higher tax rate and likewise, a worker from the second lowest ability group 2 is prone to begin to reduce human capital accumulation. Equilibrium lump-sum transfers of tax revenues form larger shares of prospective labor incomes of low-income workers, so that the tax-and-transfer system more adversely affects labor supplies of lower-income workers.

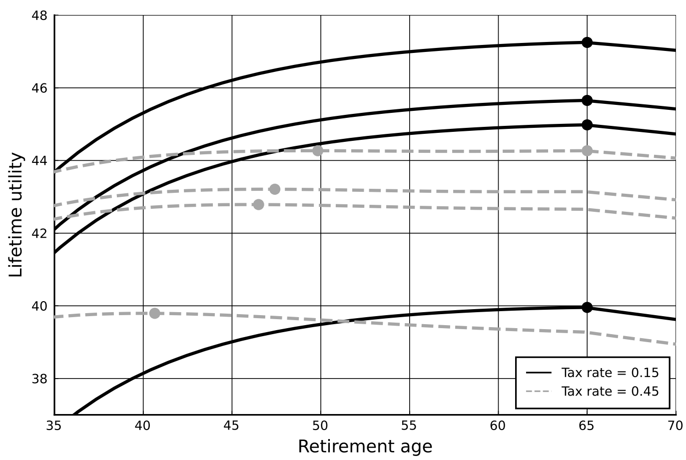


Figure 8: Lifetime utility of high school workers by ability group, conditional on working until the age on the horizontal axis, while choosing consumption and human capital accumulation optimally. For each ability group, the solid (dashed) line is computed using prices and lump-sum transfer from the tax-and-transfer equilibrium with labor tax rate $\tau_l = 0.15$ ($\tau_l = 0.45$). From top to bottom, the ability groups appear in the order 3, 4, 2, and 1. A bullet shows an optimum, which is unique on each line except in the case of ability group 3 and $\tau_l = 0.45$ when there are two optima (on the top dashed line).

When high school workers accumulate less human capital in Figure 7, they also shorten their career lengths as indicated by the declining average retirement age of high school workers along the light solid line in Figure 2(b). Except for those from ability group 3, this shortening of career lengths in response to higher labor income taxation happens gradually for high school workers. In equilibria over the tax range 0.42–0.52, high school workers of ability group 3, are indifferent between two starkly different career strategies: working until the official retirement age 65 with high end-of-life human capital versus retiring much earlier with little human capital accumulation on the job. To shed light on these outcomes, Figure 8 shows lifetime utilities of a high school worker, conditional on working continuously until various retirement ages, while choosing consumption and human capital accumulation optimally. For each ability group, a solid line refers to the baseline equilibrium with labor tax rate $\tau_l = 0.15$ and a dashed line refers to an equilibrium with tax rate $\tau_l = 0.45$. Since high school workers from ability group 3 are the most productive ones and hence, the top solid and dashed line describes ability group 3, while ability groups 4, 2 and 1 lie below in that descending order.

In the baseline equilibrium, the lifetime utility of high school workers is strictly concave in

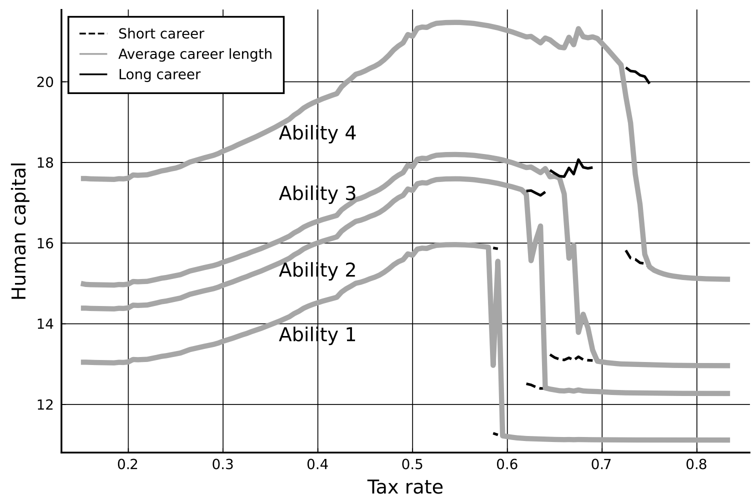


Figure 9: End-of-life human capital of college workers by ability group as a function of the labor tax rate. As in Figure 7, a dark solid and dark dashed line for an ability group indicate that workers are indifferent between accumulating relatively high and low human capital, respectively, whereas the light solid line inbetween shows a worker's average human capital in an equilibrium. Tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as equal lump-sum transfers to agents.

retirement age and all optima occur at a kink at the official retirement age 65, as marked by a bullet on each solid line in Figure 8. In the equilibrium with $\tau_l = 0.45$, a non-concavity in the lifetime utility for ability group 3 results in two optima: at the official retirement age 65 and early retirement at age 50, respectively, as marked by two bullets on the top dashed line in Figure 8. The former (latter) optimum is associated with high (low) end-of-life human capital, as shown on the dark solid (dashed) line for ability group 3 in Figure 7. While workers are indifferent between such two career strategies at tax rates in the range 0.42–0.52, market clearing pins down equilibrium *fractions* of high school workers in ability group 3 who adopt the two strategies, resulting in an average end-of-life human capital depicted by the light solid line for ability group 3 in Figure 7.

For college workers of different ability groups, Figures 9 and 10 depict end-of-life human capital and retirement age, respectively, as functions of the labor tax rate. Qualitatively, outcomes resemble those for high school workers in ability group 3. First, a college worker's end-of-life human capital initially increases with increases in the tax rate. As for a high school worker in ability group 3, the explanation is that these workers' higher supplies offset lower supplies of college human capital by other workers. But unlike high school workers in

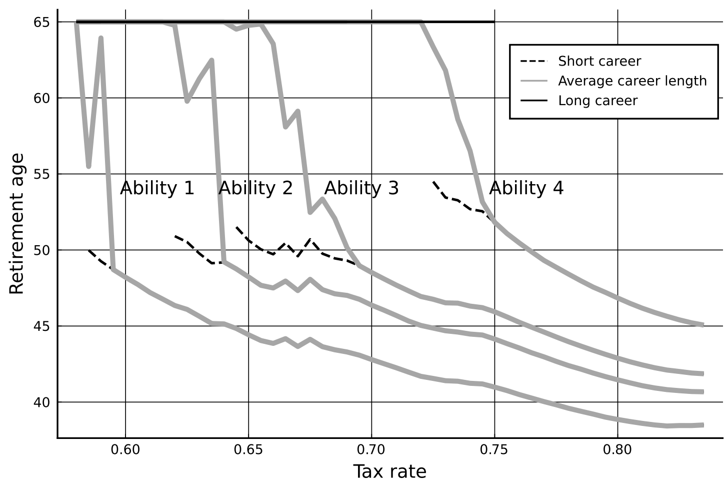


Figure 10: Retirement age of college workers by ability group as a function of the labor tax rate. Farthest to the left, all workers retire at the official retirement age 65. Next, for each ability group, there is a range of taxes over which workers are indifferent between retiring at age 65 and at an earlier age on the dashed line, whereas the light solid line shows the average retirement age in an equilibrium. Thus, when a light solid line merges with the dashed line, all workers of that ability group retire early.

ability group 3 who make up for other high school workers' shortening their career lengths and investing less in human capital, college workers' responses initially make up only for a falling college enrollment rate. Specifically, at tax rates below 0.62, practically all college graduates continue to work until the official retirement age 65 while the college enrollment rate falls monotonically in the tax rate, as shown by the dark solid line in Figures 2(b) and 3(a), respectively.[18] Second, like high school workers in ability group 3, college graduates of each ability group eventually enter a tax range over which they are indifferent between two very different career strategies, as can be inferred from Figures 9 and 10.

Nonconvexities in the space of career strategies arise not only in the baseline economy

[18]College graduates from the lowest ability group 1 do reduce their retirement age and human capital accumulation just before τ_l reaches 0.60, as shown in Figures 9 and 10. However, at those tax rates, less than 0.1 percent of agents in ability group 1 attend college and hence, their effect on aggregate outcomes is small. Furthermore, this explains why there can be large swings in the fractions of these workers who choose long and short career lengths, respectively, during the transition from all of them choosing the long career length up and until tax rate $\tau_l = 0.58$ to everyone choosing the short career length at and beyond $\tau_l = 0.595$ in Figure 10, without causing any noticeable equilibrium repercussions on choices of career lengths and human capitals of other worker categories. Presumably, these swings are spurious outcomes as the computer algorithm seeks to divide an almost nonexistent category of workers, college graduates in ability group 1, across two starkly different career choices between which they are indifferent over a small tax interval.

when workers might choose a corner solution to retire at the official retirement age because of an implicit extra tax wedge on labor income after that age. They can also arise under the social security reform when all workers are at interior solutions with respect to career length as well as in a laissez-faire economy without any government intervention. The Ben-Porath human capital technology is the source of the nonconvexities. Gains from investments in human capital can only be reaped by choosing to work for long periods of time. If government policies or the human capital technology itself becomes less conducive to human capital accumulation, an individual agent can find himself choosing between a long labor market career with considerable investment in human capital or a much shorter career length with little or no investment in human capital. At high enough tax rates, all workers in the baseline economy choose not to accumulate human capital on the job so that end-of-life human capital in Figures 7 and 9 eventually equal initial endowments of human capitals for all ability groups and schooling choices.

Despite discontinuities in an individual agent's labor market choices, the equilibrium ratio of aggregate college to high school human captial in Figure 3(b) is relatively smooth, and so are average retirement ages of college and high school workers in Figure 2(b). In a model with more heterogeneities across agents, the smoothing could occur as increases in the tax rate would gradually induce more agents to cross the threshold of making such discontinuous changes in labor market choices; in our model with limited heterogeneity, i.e., only 8 combinations of ability levels and schooling choices, the convexification of aggregate outcomes arises instead because of subsets of agents become indifferent between two very different labor market strategies, so that as the tax rate increases market clearing gradually adjusts equilibrium fractions of workers who adopt different strategies. The college enrollment rate in the economy with lump-sum rebates of tax revenues is an aggregate outcome that does not show smooth dependence on the tax rate. Specifically, the solid line in Figure 3(a) levels out over the tax range 0.62–0.75 as college workers gradually transition from retiring at the official retirement age 65 to retiring earlier in Figure 10. What happens here is that an underlying monotone decline in the absolute aggregate quantity of college human capital continues uninterruptedly though, over the tax range 0.62–0.75, this decline no longer manifests as a falling college enrollment rate but as college workers choosing shorter career lengths and less human capital investment on the job. When a tax rate equal to or exceeding 0.75 has induced all workers to retire early, the college enrollment rate resumes its decline and does so precipitously.[19]

[19] Appendix C analyzes aggregate outcomes in a small open economy version of our model.

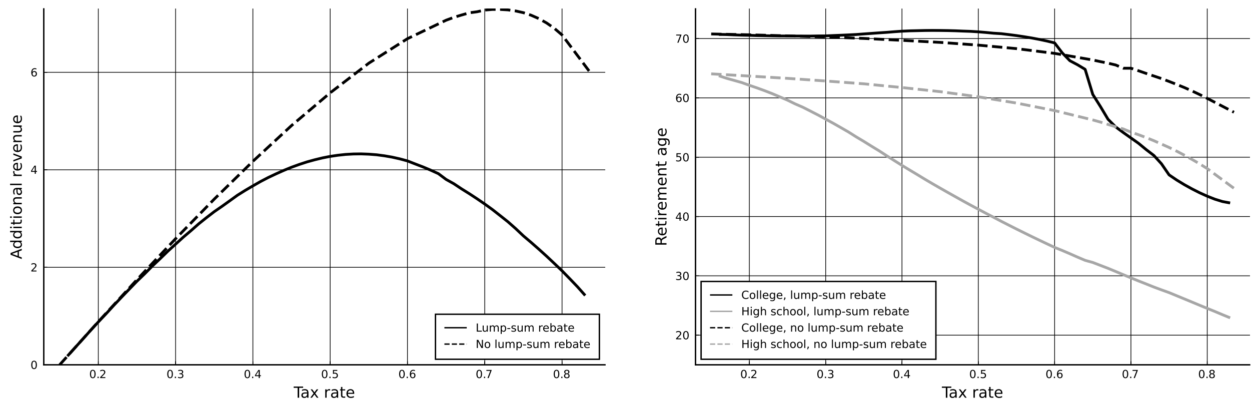


Figure 11: Laffer curves and retirement ages under the social security reform, as a function of the labor tax rate, and whether or not tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as equal lump-sum transfers to agents. In Panel A, the solid and the dashed line is the Laffer curve with and without lump-sum rebates, respectively. In Panel B, the dark and light solid (dashed) line is the average retirement age of college and high school workers, respectively, with (without) lump-sum rebates.

6 Taxation under social security reform

In the tax experiment of Section 5 in an economy under the social security reform of Section 4, we found that the Laffer curves in Figure 11(a) closely resemble those for the economy under the old social security system in Figure 2(a). But there are differences in retirement age outcomes. Recall from the second column of Table 5 that the social security reform induces college workers to retire *later*, on average at age 70.8, while high school workers choose to retire *earlier* at age 63.9. These two outcomes appear at the farthest left points of the dark and light lines, respectively, in Figure 11(b), i.e., at the baseline tax rate $\tau_l = 0.15$.

Under the social security reform, agents can no longer be stuck at a corner solution of retiring at the official retirement age 65. There are two telltale differences between the economy under the old social security system in Figure 2(b) and the economy under the social security reform in Figure 11(b) when tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as lump-sum transfers (solid lines). In the former economy, for small increases in taxation above the baseline rate $\tau_l = 0.15$ high school workers initially remain stuck at the official retirement age. But in the latter economy, being at interior solutions with respect to career lengths, high school workers' average retirement age decreases with the first increments in the tax rate. In the former economy, until fairly high levels of taxation college workers increase their lifetime supplies of human capital by accumulating more human capital on the job while they are still stuck at the official retirement age. By way of contrast in the latter economy, because they are at interior solutions with respect to their choices of career lengths, college workers also increase their average retirement age over a mid-range of tax rates in Figure 11(b).

Eventually, the tax rate becomes so high that all workers, both in the economy under the old social security system and in the economy under the social security reform, choose to retire before the official retirement age; after that equilibrium outcomes are the same in the two economies. That critical point occurs when the tax rate rises above $\tau_l = 0.74$ ($\tau_l = 0.72$) in the case of (no) lump-sum rebates of tax revenues, so that the associated lines in Figures 2 and 11 are identically the same beyond that tax rate. For this range of taxes rates in the economy under the old social security system, the implicit extra tax wedge on labor income after age 65 becomes irrelevant and workers' choices of career lengths are the same as they would be if they were living in an economy under the social security reform without the extra tax wedge.

Despite these differences, if we were to conceal the numbers on the vertical axis in Figures 2(b) and 11(b), profiles of college workers' average retirement age as functions of the tax rate would seem to be very similar. While an official retirement age at 65 "anchors" career lengths of college workers in Figure 2(b), something else apparently anchors their career length in Figure 11(b). That "something else" is a combination of HLT's earnings dynamics of college workers relative to high school workers, which in our model causes college workers to choose longer careers, and our configuration of the fixed disutility B of working and the human capital depreciation schedule in Figure 1 (solid line). These model features induce college workers to retire at a steep slope with respect to the human capital depreciation schedule. Consequently, a "wedge" arising from old-age human capital depreciation requires relatively strong forces to dislodge college workers' choices of career length under the social security reform, producing effects similar to those induced by the implicit extra tax wedge on labor income after the official retirement age under the old social security system.

7 Aggregate labor supply elasticity

In view of agents' dichotomous choices of whether to work full-time or not at all in a period, we let an employment-to-population ratio Ω represent aggregate labor supply, as depicted in Figure 12(a) for the tax experiments in Section 5. To measure sensitivity of Ω to labor taxation, we compute the aggregate labor supply elasticity Ξ with respect to the net-of-tax

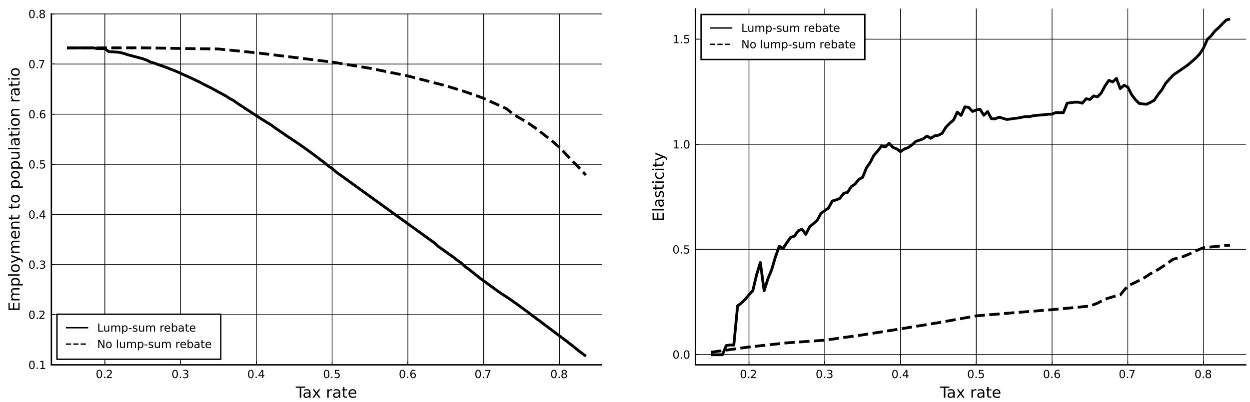


Figure 12: Employment to population ratio (Panel A) and aggregate labor supply elasticity (Panel B) as a function of the labor tax rate, and whether or not tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as equal lump-sum transfers to agents. In both panels, the solid and the dashed line displays outcomes with and without lump-sum rebates, respectively.

rate $1 - \tau_l$:

$$\Xi = \frac{\partial \Omega}{\partial (1 - \tau_l)} \frac{1 - \tau_l}{\Omega}, \tag{23}$$

as reported in Figure 12(b).[20] The solid line in the two panels of Figure 12 pertains to the economy in which tax revenues from labor taxation above the baseline rate $\tau_l = 0.15$ are handed back as equal lump-sum transfers to agents, while the dashed line describes the economy without such lump-sum rebates.

Consistent with outcomes described in Section 5, the aggregate labor supply elasticity is much subdued in the economy without lump-sum rebates. After a slow, gradual increase, there is a significant uptick in the elasticity first when the equilibrium interest rate in Figure 4(a) (dashed line) begins to increase exponentially around tax rate $\tau_l = 0.7$ because of an ever greater scarcity of physical capital. Labor supply is more severely distorted in the economy with lump-sum rebates. After a brief initial range of a zero aggregate labor supply elasticity because all workers are stuck at a corner solution of retiring at the official retirement age 65, the elasticity quickly increases above one where it levels off around 1.2 until increasing again after tax rate $\tau_l = 0.7$.[21]

[20]Because of small, unnoticable fluctuations in the employment-to-population ratio Ω in Figure 12(a), we use a 7-point centered moving average of Ω to smooth the computations of the elasticity Ξ in Figure 12(b). Instead of truncating the elasticity calculations at the end points of the tax range, we use fewer points to form the moving averages of Ω around the end points.

[21]When preferences are the same as our utility function (4) in the framework of Ljungqvist and Sargent

The rather flat aggregate labor supply elasticity of around 1.2 in the economy with lump-sum rebates might seem surprising since college workers are stuck at the corner solution of retiring at the official retirement age 65 practically up and until tax rate $\tau_l = 0.6$. But decreases in college enrollment rates at lower tax rates in Figure 3(a) (solid line) mean that agents who would be college workers at lower tax rates have now chosen to be high school workers and to retire at the earlier ages chosen by high school workers.

If we compare outcomes in Figure 12(a) with corresponding employment-to-population ratios in the Section 6 economy under social security reform, the Section 6 economy differs only in having a labor supply that is initially slightly elevated. So aggregate labor supply elasticities are quite similar in the two economies.

8 Sensitivity and numerical issues

8.1 Disutility of working and human capital depreciation

If workers choose a corner solution of retiring at the official retirement age $\eta_p = 65$, the fixed disutility B of working and the two parameters of the logistic function for efficiency units of human capital, ϕ_1 and ϕ_2, are not well pinned down under our initial social security system. Furthermore, in countries like the U.S. that have moved from earlier social security arrangements that resembled our initial social security system to reformed arrangements that more nearly resemble our stylized social security reform, it is possible that transitions are incomplete. For these and related reasons, rather than seeking to infer these parameters from U.S. data, we confine our inquiry to a sensitivity analysis of the fixed disutility of working and the decline of efficiency units of human capital in old age.

Based on HLT's and our common assumption of an 80-year lifespan, we set an inflection point $\phi_2 = 75$ for the Section 3.4 logistic function that describes efficiency units of human capital, so that efficiency units of human capital will have declined by 50 percent at the age of 75. That makes all agents want to retire during their lifetimes under plausible parameterizations of the fixed disutility of working. While holding $\phi_2 = 75$ constant, we focus here on alternative configurations of the fixed disutility B of working and the "slope" coefficient ϕ_1

(2014, p. 9, eq. (26)), the corresponding aggregate labor supply elasticity is analytically equal to one, regardless of the exponent on past work experience in a power function that determines the current wage in the learning-by-doing technology. We obtain a similar elasticity of around 1.2 over a substantial middle range of tax rates in a time-averaging version of HLT's growth model with a Ben-Porath human capital technology and several other features that differ from the Ljungqvist-Sargent framework.

of the logistic function, keeping in mind that our calibration target is that all agents choose to retire at the official retirement age 65 in the baseline steady state.

Given our baseline parameterization of the social security system and $\phi_2 = 75$, we investigate the first Section 3.4 calibration step in the following way. For each value of the fixed disutility B of working, we compute a range of slope coefficients ϕ_1 of the logistic function that attain our calibration target that all agents choose to retire at the official retirement age 65. As we perturb toward either a lower or a higher value of B, the width of the range of permissible values of ϕ_1 eventually starts to shrink, and at some point becomes zero because the value of the fixed disutility B of working has become so low or so high that no value of ϕ_1 lets us attain the calibration target. Instead, for any value of ϕ_1, there will be at least one category of workers, defined by their ability group and schooling choice, who want to retire either later or earlier than the official retirement age 65.

Searching over (B, ϕ_1) in this way yields approximate end-point coordinates of $(0.59, 0.09)$ and $(0.9, 0.31)$, respectively. Our baseline parameterization was actually guided by outcomes of this search process when we selected an intermediate pair (B, ϕ_1) subject to the additional constraint that a noticeable decline in efficiency units of human capital should not occur until workers are in their 60's. Accordingly, in Section 3.4 we picked $\phi_1 = 0.2$ with the corresponding age profile of the conversion of one unit of human capital into efficiency units depicted in Figure 1, along with profiles for the two 'end-point' values of $\phi_1 = 0.09$ and $\phi_1 = 0.31$, respectively. Associated with parameter value $\phi_1 = 0.2$ is a range of the fixed disutility of working, $B \in (0.77, 0.84)$, that makes all agents choose to retire at age 65; we selected the midpoint $B = 0.8$ for our baseline parameterization. We now examine how our analyses would have changed if we had instead adopted one of the two pairs of end-point coordinates for the parameter pair (B, ϕ_1).

Appendix D, Tables D.1 and D.2 report steady states that we would have attained if we had let parameters (B, ϕ_1) be 'low' at $(0.59, 0.09)$ or 'high' at $(0.9, 0.31)$, while retaining the other calibration steps in Section 3.4. These alternative calibrations change little overall.[22] For alternative initial steady states and future steady states after skill-biased technological change, outcomes are very similar to those of our parameterization in the last two columns of Table 3 because agents continue to retire at the official retirement age 65 with one exception:

[22] The only substantive difference is that the aggregate stock of physical capital for 'low' (B, ϕ_1) increases much more in response to skill-biased technological change as compared to the other two economies. The explanation has to do with the coefficient $\phi_1 = 0.09$ that causes efficiency units of human capital to decline significantly earlier in life as shown in Figure 1, which creates a larger demand for lifecycle savings among workers whose labor earnings taper off much earlier in life.

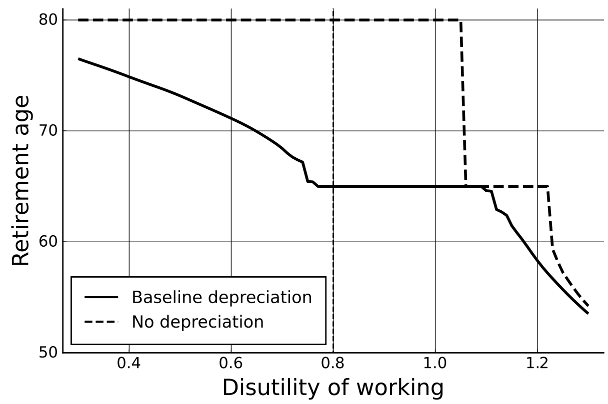


Figure 13: Retirement age of a single agent in the baseline economy as a function of her "perturbed" disutility of working, $\hat{B}$. The individual is either a high school worker of ability group 1 (Panel A) or a college worker of ability group 4 (Panel B), who is subject to either baseline depreciation of human capital (solid line) or no depreciation (dashed line).

with 'high' (B, ϕ_1), high school workers of the lowest ability group 1 retire a couple of years earlier in the steady state after skill-biased technological change. This equilibrium change foreshadows outcomes under social security reform. Thus, as compared to our steady state after social security reform in the second column of Table 5, all high school workers retire 1-3 years *earlier* with 'high' (B, ϕ_1), while they retire 1.5-2 years *later* with 'low' (B, ϕ_1). In both cases, college workers retire about one year later than in our steady state after social security reform.

To offer perspective on effects of different values of the fixed disutility B of working, consider a single worker who enters our baseline economy with his own value of $\hat{B} \in [0.3, 1.3]$. Figure 13(a) shows the optimal retirement age as a function of $\hat{B}$ for a single high school worker of the lowest ability group 1 that contributes most to the pool of high school workers. The solid and dashed lines depict retirement ages conditional on that agent having the baseline logistic function for the efficiency units of human capital and instead of experiencing no depreciation, respectively. Figure 13(b) is the corresponding graph for a single college worker of the highest ability group 4 that contributes the most to the pool of college workers. Since we consider a single individual who enters the economy with these different characteristics, that worker's arrival bring no effects on aggregate outcomes, so that the interest rate and prices of human capital stay at their baseline steady state values.

At sufficiently low values of $\hat{B}$, all of the individuals in Figure 13 work beyond the official retirement age 65 and, if there is no depreciation of human capital, both high school and college graduates work their entire lifespans until age $\overline{\eta} = 80$. As we gradually increase the fixed disutility of work, over some interval for $\hat{B}$ agents eventually get stuck at the corner and retire at age 65. At the baseline parameter value $\hat{B} = B = 0.8$, points on a solid line depict a representative high school worker of ability group 1 and a representative college worker of ability group 4, respectively, in our baseline economy. For reasons explained in Sections 4 and 5, a high school worker is more likely than a college worker to shorten his career length. This is reflected here in a representative high school worker of ability group 1 being closer to the right end of the interval of $\hat{B}$ over which it is optimal to retire at age 65 than is a representative college worker of ability group 4, i.e., it takes a smaller increase in $\hat{B}$ to induce the high school worker in Figure 13(a) to retire earlier than age 65 as compared to the college worker in Figure 13(b). Turning to corresponding flat segments of the dashed lines for the two individuals whose human capitals don't depreciate, there is no overlap of the two segments, so there is no common value of $\hat{B}$ that would make both the high school and the college worker choose to retire at the official retirement age 65. Appendix D, Table D.3 shows that this outcome carries over to general equilibrium. In particular, after setting $B = 1$ and shutting down depreciation of human capital by setting $e(n) = 1$ for all ages n, we retain the rest of our baseline parameterization except for recalibrating the fraction of capital held by agents, κ, so that we hit HLT's and our common target for the initial steady-state interest rate. In the resulting steady state, all high-school workers retire before the official retirement age 65, while college workers of the first three ability groups retire at age 65 and those of the highest ability group 4 never retire. Next, we confirm that neither lowering nor raising B in such a perturbation of the steady state can induce all agents to retire at age 65. Thus, human capital depreciation is essential for arranging parameter values for our baseline steady state.

Sharp drops in retirement ages in Figure 13 and accompanying sharp drops in the end-of-life human capitals stock (not shown) are manifestations of Section 5.2 nonconvexities in the space of career strategies, but now they show up in responses to perturbations of an agent's disutility $\hat{B}$ of work instead of to our earlier perturbations of a tax rate. Effects are especially striking when human capital does not depreciate (the dashed lines) and an agent switches from working his entire life and instead retires at the official retirement age 65. Besides intrinsic features of the Ben-Porath human capital technology that can put nonconvexities into the problem of choosing a career strategy, an additional factor affects transitions from working an entire life-time to retiring at age 65. Thus, under the initial social security system, someone who works beyond age 65 sacrifices the present value of his foregone social security benefits, so that an implicit extra tax wedge that induces all workers in the baseline economy to retire at age 65 now also awards a large capital gain to anyone who retires at age 65 instead of working until 80: by making that downward jump in retirement age, the worker in one swoop can collect the present value of social security benefits between ages 65 and 80. In contrast, for agents who confront the baseline depreciation of human capital in Figure 13 (solid lines), the lure of that capital gain is diminished by an improved conversion of human capital into efficiency units as an agent "walks up" along the schedule in Figure 1 (solid line) so that the decreases in the retirement age toward the official retirement age becomes more gradual. Eventually, as optimal retirement ages continue to decline at the far right ends of Figure 13, the solid and dashed lines for a worker converge so that it no longer matters whether human capital depreciates since the rate of conversion of human capital into efficiency units approaches being one-to-one at the chosen retirement age.

8.2 Lumpiness in labor supply

Following HLT, our model period is one year, maximum life length is 80 years, and career length is the number of periods an agent works. This configuration creates lumpiness in an individual's labor supply. Limited heterogeneity among employed workers – four ability groups and two schooling levels – means that lumpiness can also occur in aggregate labor supply outcomes. That potentially complicates computing an equilibrium. Thus, for the Section 5 tax-and-transfer scheme our algorithm failed to converge to an equilibrium for some tax rates because choices of a significant group of agents between retiring in one year versus a year earlier could not be reconciled with an equilibrium value for the lump-sum transfer. On one hand, they *would* like to retire a year earlier if they could receive the *larger* lump-sum transfer that would be the equilibrium amount if their entire group did *not* retire a year earlier. On the other hand, they *would not* like to retire a year earlier if they were to receive the *smaller* lump-sum transfer that would be the equilibrium amount if the group as a whole had retired a year earlier. While the problem seems to resemble the Section 5.2 computational challenge that we resolved by randomizing, a key difference is that the earlier "tipping points" at which identical agents were indifferent between starkly different lifetime career choices could not have been remedied by simply using a finer model period. However, the present computational challenge is amenable to such an approach.

Before turning to such a solution that involves relaxing the assumption that an agent has to devote an integer of years to work, we emphasize that individual labor supply lumpiness is at the heart of time averaging. A high labor supply elasticity arises from decisions taken at an extensive margin of whether or not to work in a given discrete period, coupled with assumptions on preferences and technologies that impart lumpiness to those decisions. Thus, on the preference side, a decision to work positive hours can come from a fixed disutility of work; and on the technology side, nonconvexities can favor full-time over part-time work that would come from a lower productivity and hence a lower hourly wage for part-time work. Nevertheless, it seems implausible to require that such lumpiness of labor supply applies to an entire *year* as it does with our annual model period. Consequently, to allow for someone to work just a fraction of the last year of a career, we now proceed under the assumption that the disutility B of work is scaled by the fraction of a year devoted to work.

As described in Appendix B.3.1, we use cubic splines to interpolate between the computed points that comprise a curve like those depicted in Figure 8. Our purpose is to approximate an agent's maximal lifetime utility while offering a continuous value of a career length by allowing an agent to work just a fraction of his retirement year (but not earlier years). Next, after approximating an optimal continuous retirement age with that interpolated curve, we linearly interpolate an agent's lifetime paths of human capital, asset holdings, and consumption. This is a tractable way to approximate finer sub-periods with respect to an agent's choice of career length while still retaining HLT's annual model period.

For agents who find it optimal not to accumulate human capital on the job, there is an analytical formula for an optimal continuous retirement age that comes directly from the solution to the annual problem. It is exact in the sense that it gives the same retirement age that we would have obtained if we had assumed that career length were a continuous choice variable. Equivalence of a two-stage optimization problem and a single maximization problem sheds light on the workings of our model and its HLT building blocks. First, HLT's assumptions of complete markets and CES preferences over consumption imply that the shape of a consumption profile from the first stage will also be optimal in a second stage. The consumption profile just shifts proportionately down or up depending on whether an optimal continuous retirement age is shorter or longer than the annual retirement age from the first stage. Second, it is easy to show that an optimal continuous retirement age falls in the open interval shifted down or up one year relative to the optimal annual retirement age from the first stage. (The interval is open since its endpoints have already been shown to be suboptimal in the first stage.) As detailed in Appendix B.3.1, the optimal continuous retirement age is then pinned down by the first-order condition with respect to an incremental change in career length around an optimal annual retirement age from the first stage.[23]

[23]In contrast to the spline interpolation method, the analytical formula applies the proper present-value

We can use the analytical formula to check for the accuracy of the spline interpolation method. Under the Section 5 tax-and-transfer scheme, at high enough tax rates every agent eventually chooses not to accumulate human capital on the job. For example, at tax rate $\tau_l = 0.61$, high school workers of all four ability groups are accumulating virtually no human capital on the job in Figure 7. Thus, the analytical formula provides exact optimal continuous retirement ages for these workers that we compare to spline interpolations in Appendix B.3.1, Table B.1. Specifically, using equilibrium prices and lump-sum transfers computed from our model with discrete annual retirement ages, we calculate continuous retirement ages with both the spline interpolation method and the analytical formula. The spline interpolations yield approximations very close to the exact continuous retirement ages of the analytical formula, the largest deviation being only half a month. This finding makes us confident to use the spline interpolation method, whereas the analytical formula would not apply for the general case of agents choosing to accumulate additional human capital on the job.

9 Concluding remarks

Our inclusion of endogenous career lengths and indivisible labor substantially enriches HLT's general equilibrium analysis of labor market outcomes with respect to schooling choices and on-the-job human capital accumulation. As a starting point, we show how the addition of a social security system can rationalize HLT's entire analysis when career lengths are endogenous. If the retirement system imposes an implicit extra tax wedge on labor income after an official retirement age in the form of lost social security benefits, agents who are heterogeneous in abilities and schooling levels might nevertheless choose corner solutions at an official retirement age.

HLT's labor force composition of college and high school graduates highlights earnings profiles as determinants of career length. In a learning-by-doing setup, Ljungqvist and Sargent (2014, sec. 3) show that the more elastic is an earnings profile to accumulated working time, the longer is a worker's career. We have shown how this outcome extends to HLT's Ben-Porath human capital technology where the technology of college workers is estimated to be more productive and hence, college workers are prone to choose longer careers than high school workers. Longer careers of college workers emerge from under a

discount factor associated with the 'annual windows' in which any of the repercussions of the optimal perturbation falls, including changes in consumption over the entire lifespan and, in particular, the marginal change in career length and disutility of working where the proper annual discount factor depends on if the optimal perturbation calls for a shorter or a longer career length.

social security reform that removes the implicit extra tax wedge on earnings after the official retirement age. Because we assume preferences that are consistent with balanced growth, it is not the level of earnings that matter but only terms on which more human capital can be accumulated on a job.

In a companion paper, Heckman, Lochner, and Taber (1998b) compared partial- and general-equilibrium effects of tax reforms that favor investments in human capital. Given a fixed interest rate and skill prices, while college enrollment and on-the-job human capital accumulation can increase markedly in a partial equilibrium, responses are much diminished in a general equilibrium in which a higher interest rate emerges from a relative scarcity of physical capital. That interest rate channel is very much present in our tax experiments, while general-equilibrium dynamics coming from endogenous career lengths further augment it. When tax revenues are handed back as lump-sum transfers, changes in the labor tax rate bring stark interdependencies across ability and schooling levels. Until fairly high tax rates on labor income, the negative impact on college human capital comes from falling college enrollment, while those who do graduate from college continue to work long careers and to invest in human capital on the job. In contrast, high school workers shorten career lengths and reduce investments in human capital, especially those in the lowest ability groups. Thus, those high labor tax rates exacerbate a labor market bifurcation between highly active workers and some who have been sidelined or marginalized.

We encountered a perhaps neglected property of a Ben-Porath human capital technology in the form of nonconvexities in choice sets for career strategies. For investments in human capital to pay off, an investor needs to choose a long enough career. That complicates responses to changes in the disutility of working, productivity of the human capital technology, and tax rates. We discovered that the optimal response to an incremental change in determinants can be discontinuous and possibly even provoke large shifts in career length and human capital investments. For example, in our tax experiments with lump-sum rebates of tax revenues, such big responses are made by the most able high school workers and college workers of all ability groups who will, at some incremental tax increase experience a “tipping point” that takes the form of a discrete change from a long career with substantial human capital investments to a significantly shorter career with much less on-the-job human capital accumulation.

Thus, our time-averaging version of the HLT model brings out challenges for economic policy making. Some of our social security reforms and tax-and-transfer arrangements increase the hazard of a “dual labor market” marked by different labor market participation rates across schooling and ability groups. Agents with steeper earnings profiles, like college workers, are prone to participate more robustly in the labor market. The steepness of an agent's earnings profile depends on both his Ben-Porath technology and his choice of how much human capital to accumulate on the job. This affects the efficacy of tax-and-transfer schemes. If government policies eventually bring high enough tax wedges and other distortions, then even those agents having the most productive human capital technologies who have historically supplied more labor more robustly become a source of labor market fragility and might dramatically shift their career strategies toward earlier retirements and less human capital accumulation. The prospect of an implosion of economic activity raises the stakes for avoiding policies that erode labor market participation among substantial proportions of the labor force.

References

Auerbach, Alan J. and Laurence J. Kotlikoff. 1987. *Dynamic Fiscal Policy*. Cambridge, MA: Cambridge University Press.

Browning, Martin, Lars Peter Hansen, and James Heckman. 1999. Micro Data and General Equilibrium Models. In *Handbook of Macroeconomics, Vol. 1A*, edited by John Taylor and Michael Woodford. Amsterdam: North Holland.

Cameron, Stephen V. and James J. Heckman. 1998. Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males. *Journal of Political Economy* 106 (2):262–333.

Chang, Yongsung and Sun-Bin Kim. 2006. From Individual to Aggregate Labor Supply: A Quantitative Analysis Based on a Heterogeneous Agent Macroeconomy. *International Economic Review* 47:1–27.

De Nardi, Mariacristina. 2015. Quantitative Models of Wealth Inequality: A Survey. Working Paper 21106, National Bureau of Economic Research.

Diamond, Peter A. 1965. National Debt in a Neoclassical Growth Model. *American Economic Review* 55:1126–1150.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998a. Explaining Rising Wage

Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents. *Review of Economic Dynamics* 1 (1):1–58.

———. 1998b. Tax Policy and Human Capital Formation. *American Economic Review* 88 (2):293–297.

Ljungqvist, Lars and Thomas J. Sargent. 2006. Do Taxes Explain European Employment? Indivisible Labor, Human Capital, Lotteries, and Savings. In *NBER Macroeconomics Annual 2006*, edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford. Cambridge, Mass.: MIT Press.

———. 2011. A Labor Supply Elasticity Accord? *American Economic Review* 101 (3):487–491.

———. 2014. Career Length: Effects of Curvature of Earnings Profiles, Earnings Shocks, Taxes, and Social Security. *Review of Economic Dynamics* 17 (1):1–20.

Pechman, Joseph A. 1987. *Federal Tax Policy*. 5th edition. Washington D.C.: Brookings Institution.

Prescott, Edward C. 2002. Prosperity and Depression. *American Economic Review* 92:1–15.

———. 2005. The Transformation of Macroeconomic Policy and Research. In *Les Prix Nobel 2004*, 370–395. Stockholm: Almqvist & Wiksell International [https://www.nobelprize.org/uploads/2018/06/prescott-lecture.pdf].

———. 2006a. Comment. In *NBER Macroeconomics Annual 2006*, edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford. Cambridge, Mass.: MIT Press.

———. 2006b. Nobel Lecture: The Transformation of Macroeconomic Policy and Research. *Journal of Political Economy* 114 (2):203–235.

Queisser, Monika and Edward Whitehouse. 2006. Comparing the Pension Promises of 30 OECD Countries. *International Social Security Review* 59 (3):49–77.

Rogerson, Richard. 1988. Indivisible Labor, Lotteries, and Equilibrium. *Journal of Monetary Economics* 21:3–16.

Samuelson, Paul A. 1958. An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money. *Journal of Political Economy* 66:467–482.

Schulz, James H. 2001. *The Economics of Aging, Seventh Edition*. Westport, Conn.: Auburn House.

Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron. 2004. Consumption and Risk Sharing Over the Life Cycle. *Journal of Monetary Economics* 51 (3):609–633.

Taber, Christopher. 2002. Tax Reform and Human Capital Accumulation: Evidence from an Empirical General Equilibrium Model of Skill Formation. *Advances in Economic Analysis and Policy* 2 (1): Article 3.

A Model Appendix

In this Appendix we begin by describing the model studied by HLT, along with how that model is calibrated. We then describe the changes that we make to the model in order to endogenize agents' retirement decisions and the recalibration of the model that this entails.

A.1 Model of HLT

In the model of HLT, individuals live from ages 19 to 80 and retire exogenously at age 65. Individuals differ in their schooling level, $S \in \{1, 2\}$ and in their type, $\theta \in \{1, 2, 3, 4\}$. At the beginning of life, individuals decide whether to enter the labor market as a high school graduate ($S = 1$) or to go to college ($S = 2$). Attending college takes four years, during which an individual is unable to work. An individual's type θ affects their initial human capital stock, ability to accumulate human capital, and their likelihood of attending college.

Individuals are endowed with one unit of time each period. During their working life, they decide how much time to spend investing in their human capital, I, and how much to spend working, $(1 - I)$. In all periods individuals decide how much to consume, C, and how much physical capital to hold in the next period K'. Conditional on a schooling option and a retirement age the recursive problem for an individual of type θ that is still working ($n \leq 65$) is as follows:

$$V_n(H, K, S, \theta) = \max_{C,I,K',H'} \frac{C^\gamma - 1}{\gamma} + \delta V_{n+1}(H', K', S, \theta) \tag{24}$$

$$\text{subject to}$$

$$K' = K(1 + (1 - \tau)r) + (1 - \tau)R^S H(1 - I) - C - X\mathbb{1}\{n = 65\} + X\mathbb{1}\{a = 18\}$$

$$H' = H + A^S(\theta)I^{\alpha_S} H^{\beta_S}$$

where r is the real interest rate, R^S is the wage for an agent of schooling level S, and τ is a tax rate, common to both capital and labor income. X is a transfer made from retiring agents to agents at the beginning of the life-cycle. For individuals that have retired ($n > 65$) the recursive problem is:

$$V_n(H, K, S, \theta) = \max_{C,K'} \frac{C^\gamma - 1}{\gamma} + \delta V_{n+1}(H, K', S, \theta) \tag{25}$$

$$\text{subject to}$$

$$K' = K(1 + (1 - \tau)r)$$

At the beginning of the life cycle, agent i chooses to attend college if:

$$(1-\tau)V^2(\theta) - D - \epsilon_i \geq (1-\tau)V^1(\theta) \tag{26}$$

where $V^S(\theta)$ is the present value of earnings for schooling choice S, D is the present value of the tuition cost of attending college and each agent draws a value of ϵ_i from a $N(\mu_\theta, \sigma)$ distribution. The mean of the college taste shock differs by an individual's type but the variance is common to all agents.

A.1.1 Equilibrium Conditions

The model is closed using the following constant returns to scale aggregate production function.

$$F(\bar{H}_1, \bar{H}_2, \bar{K}) = a_3[(1-a_2)Q^{\rho_2} + a_2\bar{K}^{\rho_2}]^{\frac{1}{\rho_2}} \tag{27}$$

where

$$Q = [a_1\bar{H}_1^{\rho_1} + (1-a_1)\bar{H}_2^{\rho_1}]^{\frac{1}{\rho_1}} \tag{28}$$

Given this production function, the skill prices are determined by derivatives:[24]

$$r = a_3a_2C\bar{K}^{\rho_2-1} \tag{29}$$

$$R^1 = a_3(1-a_2)a_1CQ^{\rho_2-\rho_1}\bar{H}_1^{\rho_1-1} \tag{30}$$

$$R^2 = a_3(1-a_2)(1-a_1)CQ^{\rho_2-\rho_1}\bar{H}_2^{\rho_1-1} \tag{31}$$

where

$$C = [(1-a_2)Q^{\rho_2} + a_2\bar{K}^{\rho_2}]^{\frac{1-\rho_2}{\rho_2}} \tag{32}$$

A.1.2 Calibration of HLT Model

The calibration procedure of HLT occurs in a number of stages. All parameters are reported in Table A.1.[25] First, a number of parameters are set exogenously: δ to 0.96, γ to 0.1, the real after-tax interest rate, $(1-\tau)r$, to 5%, R^S to to 2 for both levels of education, and the tax rate to 15%.

[24] Note, there is a typo in the skill prices reported in HLT on p15. Q should be raised to $\rho_2 - \rho_1$, rather than $\rho_2 - 1$

[25] There are a small number of parameters that HLT do not report. For those parameters, we report the values that we have used in our replication.

Table A.1: Calibration of HLT Model

Parameter	High School (S=1)	College (S=2)
$H_0^S(1)$	8.042	11.117
$H_0^S(2)$	10.063	12.271
$H_0^S(3)$	11.127	12.960
$H_0^S(4)$	10.361	15.095
$A^S(1)$	0.081	0.081
$A^S(2)$	0.085	0.082
$A^S(3)$	0.087	0.082
$A^S(4)$	0.086	0.084
α	0.945	0.939
β	0.832	0.871
δ	0.96	
γ	0.1	
μ_1	-53.02	
μ_2	-2.82	
μ_3	29.77	
μ_4	-28.65	
σ	22.41	
ρ_1	0.306	
ρ_2	-0.034	
τ	0.15	
X	30	
a_1	0.496	
a_2	0.252	
a_3	2.504	
D	3.62	

Notes: Parameters in the top section are reproduced directly from HLT. Parameters in the bottom section are those that we inferred from our replication of the model.

The human capital production function parameters, $\alpha^S, \beta^S, A^S(\theta), H_0^S(\theta)$ are estimated using NLSY data on white male earnings from 1979 to 1993, where individuals are split into four equal-sized groups on the basis of their Armed Forces Qualifying Test (AFQT) scores. Given the earnings profiles implied by these parameters, HLT calibrate μ_θ and σ to match the results of probit regressions for college attendance of each of the ability groups.

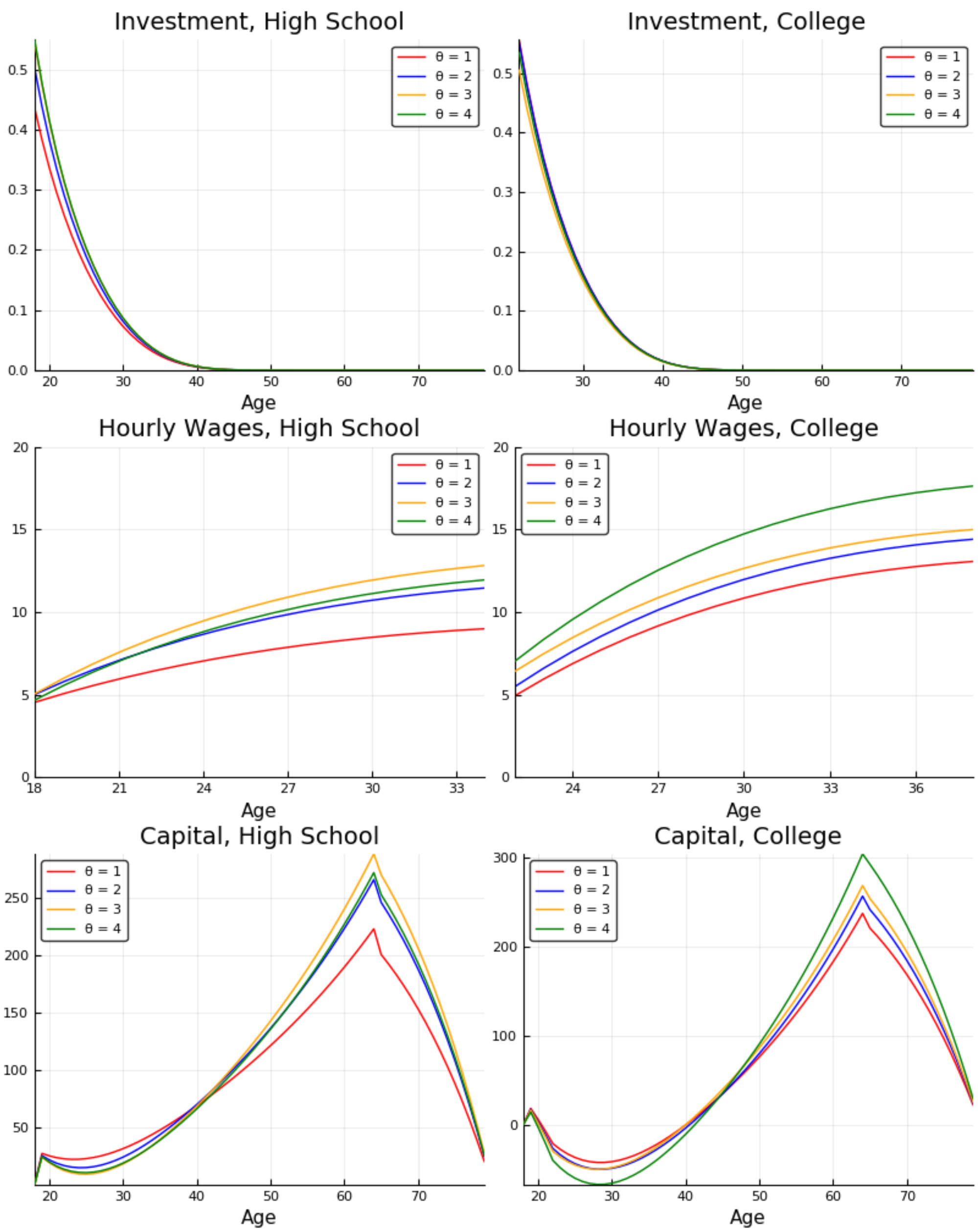
Finally, HLT use aggregate data to estimate the production function elasticities: ρ_1 and ρ_2. In the model, HLT assume a transfer of $\$30,000$ between retiring agents and agents in the first year of the life-cycle. This is chosen such that the value of the capital-output ratio in the model is 4. The values of a_1, a_2, and a_3 are then pinned down by the assumptions on r and R^S. Although HLT do not report these values directly in their paper, we include the implied values from our replication of the model in Table 1. HLT also do not report the cost of tuition in their model. However, Taber (2002) uses a very similar model in which tuition is set to 1.02 per year. Given that college takes four years, this gives us a present-value of tuition, D, equal to 3.62. Figure A.1 provides some representative plots from our replication of the HLT model.[26] In the case of human capital investment and wage profiles, these plots line up closely with Figures 2 and 3 of HLT. This is confirmed in Table A.3 where we estimate terminal values of human capital for each type of agent that are close reported in their paper. Both high school and college-educated agents initially spend around 50% of their time in human capital investment, but this rapidly declines to 0 by around age 45. Given this investment and the upward sloping profile of wages all agents save little before they are around 40. Capital holdings then increase rapidly until retirement, after which agents draw down their savings until the end of the life-cycle.

A.2 Our Model: HLT with an endogenous retirement age

Our model adheres very closely to that in HLT. The key difference is that we endogenize the agent's retirement decision: agents now choose, $\hat{n}$, the last year in which they work. In the baseline version of our model, individuals receive a pension P each period after age 65 only if they are no longer working. In Section 6 of the main paper, we modify this restriction.

[26] In our replication of HLT's model, we successfully reproduced the life-cycle paths of earnings and human capital investment, as well as the results of the SBTC experiment. However, in doing so we note that our present-value earnings of high-school graduates of different abilities are between 6.9-7.0 percent lower than those reported in Table II in HLT, and for college graduates 7.1-7.2 percent lower. As these differences are similar across schooling levels, college attendance probabilities are little affected if we calculate them using our present-value earnings or those of HLT. An implication of this is that our estimates of resulting aggregate human capital stocks are also not affected and are close to those in HLT's Figure 9. But our estimates of average lifetime earnings across schooling levels are lower than those reported in Figure 16.

Figure A.1: Life-Cycle Paths in the HLT Model



Notes: The investment and wage paths for the original steady-state were also included in the original HLT paper, in Figures 1, 2, and 3.

To ensure an interior solution for the agent's retirement decision, we assume that agents receive a disutility B for time spent either working or investing in human capital. We also assume that the productiveness of an agent's human capital is declining over time according to a function $e(n)$.

Conditional on a schooling option and a retirement age the recursive problem for an individual of type θ that is still working ($n \leq \hat{n}$) is as follows:

$$V_n^{\hat{n}}(H, K, S, \theta) = \max_{C,I,K',H'} \log(C) - B + \delta V_{n+1}^{\hat{n}}(H', K', S, \theta) \tag{33}$$

$$\text{subject to}$$

$$K' = K(1 + (1-\tau)r) + (1 - \tau - \tau_P)R^S He(n)(1 - I) - C$$

$$H' = H + A^S(\theta)I^{\alpha_S}H^{\beta_S}$$

$$e(n) = \frac{1}{1 + \exp(\phi_1(n - \phi_2))}$$

where τ_P is a payroll tax levied to fund the pension system. For individuals that have retired ($n > \hat{n}$) the recursive problem is:

$$V_n^{\hat{n}}(H, K, S, \theta) = \max_{C,K'} \log(C) + \delta V_{n+1}^{\hat{n}}(H, K', S, \theta) \tag{34}$$

$$\text{subject to}$$

$$K' = K(1 + (1-\tau)r) + (1-\tau)P(a)$$

$$P(n) = \begin{cases} 0 & \text{if } n \leq 65 \\ P & \text{if } n > 65 \end{cases}$$

As in HLT, we assume that agents that choose to attend college are unable to work for four years. We assume that during these years agents pay tuition equal to t per year. Conditional on a college attendance decision, individuals choose the retirement age that maximizes lifetime utility:

$$\hat{n}(S, \theta) = \arg\max_{\hat{n}} V_1^{\hat{n}}(H_0^S(\theta), 0, S, \theta) \tag{35}$$

Giving individuals an endogenous retirement decision and a disutility of work means that the college taste shock can no longer be denominated in present value terms. We assume that it is measured in units of utility, as in Taber (2002). Individuals choose to attend college if:

$$V_1^{\hat{n}}(H_0^2(\theta), 0, 2, \theta) + \epsilon \geq V_1^{\hat{n}}(H_0^1(\theta), 0, 1, \theta) \tag{36}$$

where $\epsilon \sim N(\mu_\theta, \sigma)$. The mean of the college taste shock differs by an individual's type but the variance is common to all agents.

A.2.1 Equilibrium Conditions

We assume that ρ_2 in HLT's original production function is equal to zero. Thus our production function is:

$$F(\bar{H}_1, \bar{H}_2, \bar{K}) = a_3 Q^{(1-a_2)} \bar{K}^{a_2} \tag{37}$$

where

$$Q = [a_1 \bar{H}_1^{\rho_1} + (1 - a_1)\bar{H}_2^{\rho_1}]^{\frac{1}{\rho_1}} \tag{38}$$

We retain a calibration of $\rho_1 = 0.306$. Rather than assuming that there is a redistribution of capital from retiring agents to the cohort being born, we instead assume that agents in our model only hold some fraction x of the aggregate capital in the economy. The remainder is held by capitalists that are outside the scope of our model.

A.2.2 Summary of Differences from HLT

As mentioned above, the key difference between our model and that in the original HLT paper is the introduction of an endogenous retirement decision. However, there are also a number of more minor differences between the two models:

1. HLT use $U(C) = \frac{C^\gamma - 1}{\gamma}$ with $\gamma = 0.1$. We use $U(C) = \log(C)$. Given their value of γ, there is little difference between these two functions.

2. We introduce a disutility of labor supply equal to B utils in each period that the agent is working.

3. We introduce the function, $e(n)$, governing the effectiveness of human capital. Note, this function only affects the ability of agents to earn labor income, and does not affect human capital accumulation. The function depends on two parameters: ϕ_1 governs the slope and ϕ_2 governs the inflection point. In Section 8 of the main paper and in Section 4 of this appendix, we discuss how sensitive the outcomes of the model are to these choices.

4. We introduce a social security system (P and τ_P). In the baseline model, agents receive a pension P each period from the age of 65, as long as they have stopped work. In the social security reform experiment of Section 6, agents receive the pension regardless of whether or not they have stopped work. This is financed by a social security tax on labor income equal to τ_P.

5. HLT have a college taste shock in present value terms. Their model involves no retirement choice or disutility of labor supply, so the optimal policy coincides with the policy that maximizes the present value of income. We replace this with a college taste shock denoted in terms of utility, as in Taber (2002).

6. HLT use a transfer scheme to maintain a capital-output ratio of 4. We remove this transfer and instead assume that agents in our economy only hold a fraction x of aggregate capital.

A.3 Calibration

We retain many of the parameters in the original HLT calibration. We use the human capital production function parameters reported in Table A.1. We also keep the production function parameter of $\rho_1 = 0.306$. As HLTs estimate of ρ_2 is not significant, we round ρ_2 to 0. Table A.2 presents the remaining parameters of our baseline calibration.

Of the remaining parameters, we set the mean of the college taste shock such that the college attendance probabilities for each type are equal to those in HLT. Given values for μ_θ, the standard deviation of the college taste shock σ governs how responsive college attendance probabilities are to changes in the skill prices R^1 and R^2. We set $\sigma = 1.5$ such that the ratio of these skill prices $\frac{R^2}{R^1}$ rises from 1 to 1.08 between the original and new steady-states in response to HLT's original experiment of skill-biased technical change.

We exogenously set ϕ_2 to 75. This implies that agents aged 75 are only 50% as productive as young agents with the same human capital. We set the slope of the efficiency function, ϕ_1, to 0.2. Given this assumption, we pick the disutility of labor supply such that all agents choose to retire at 65 in the original steady-state of the model.

We set the three remaining parameters in the production function such that the three prices are the same in the original steady-state as they are in HLT.[27] We set the value of the pension to 8, in order to generate an average replacement rate of approximately 40%.

[27]These values are very similar to those we estimated from the original HLT model, reported in Table A.1.

Figure A.2: Life-Cycle Paths in the Our Baseline Model

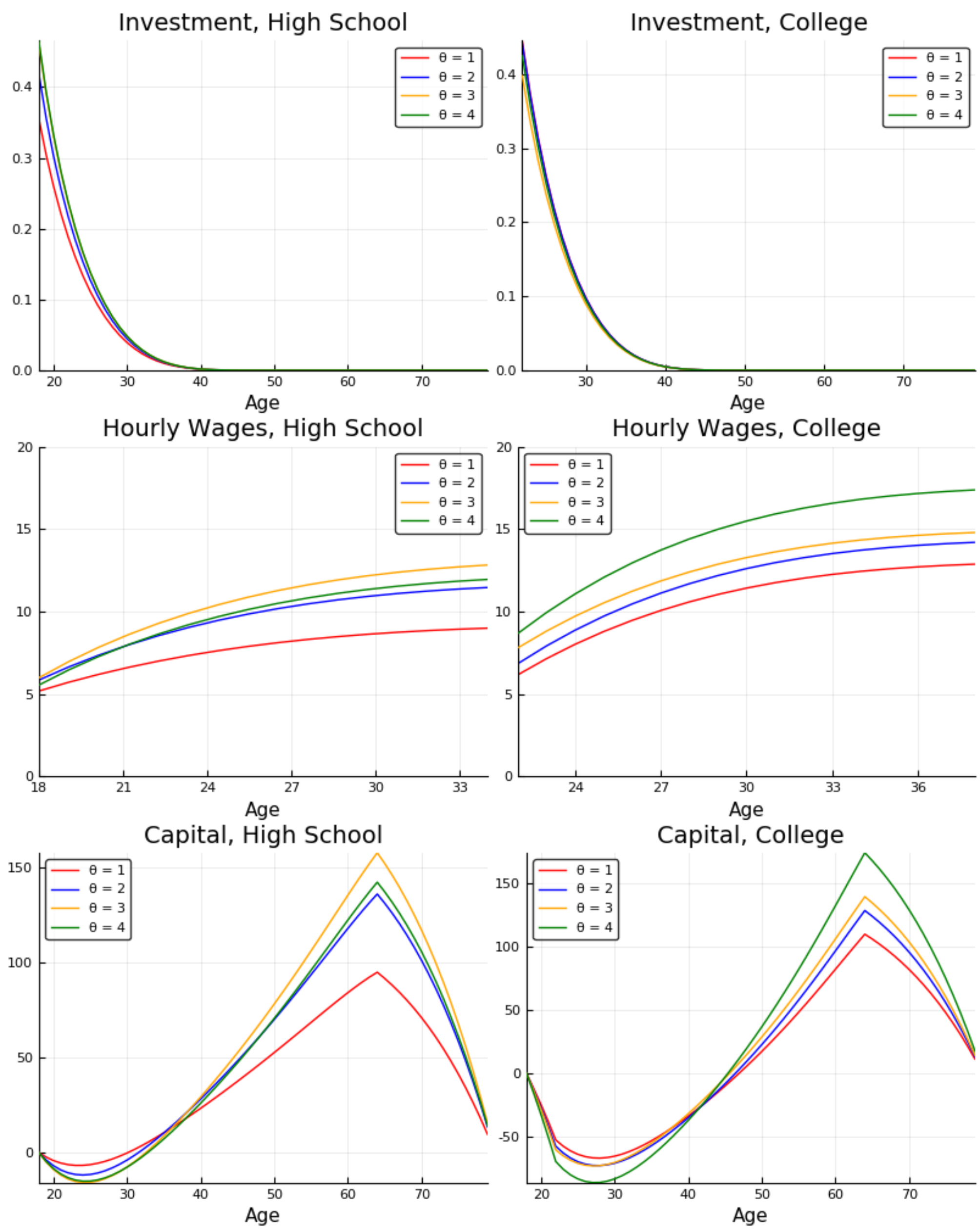


Table A.2: Our Calibration

Parameter	Value	Definition	Source/Target
μ_1	-6.819	Mean of college taste shock ($\theta = 1$)	College Attendance
μ_2	-2.808	Mean of college taste shock ($\theta = 2$)	College Attendance
μ_3	-0.912	Mean of college taste shock ($\theta = 3$)	College Attendance
μ_4	-4.587	Mean of college taste shock ($\theta = 4$)	College Attendance
σ	1.5	SD of college taste shock	SBTC Experiment
a_1	0.475	Production function parameter	$r(1-\tau_k) = 0.05$
a_2	0.235	Production function parameter	$R^1 = 2$
a_3	2.554	Production function parameter	$R^2 = 2$
ϕ_1	0.2	Slope of efficiency units	Retire at 65
ϕ_2	75	Inflection point in efficiency units	N/A
B	0.8	Disutility of labor supply	Retire at 65
τ_P	0.1	Social security tax	N/A
P	8	Pension	Replacement rate $\approx 40\%$
x	0.377	Fraction of capital held by agents	$\frac{K}{Y} = 4$

Notes: θ denotes the four ability types.

This pension is financed by a linear 10% social security tax on labor income. Finally, we assume that the agents in our model only hold 38% of the total capital in the economy. This ensures that the capital-output ratio is 4 in the original steady-state. We assume that the remaining capital is held by capitalists who are outside the scope of our model. Figure A.2 reports the life-cycle paths for human capital investment, wages, and capital holdings in the baseline calibration of our model. These are very similar to those in the HLT model shown in Figure A.1.

A.4 Original HLT Experiment: SBTC

As mentioned in the previous section, we calibrate the standard deviation of the college taste shock such that in the long-run the ratio of skill prices rises from 1 to 1.08 in response to the skill-biased technical change studied by HLT.

As in HLT, we assume that $\log\left(\frac{\alpha_1}{1-\alpha_1}\right)$ declines by 3.6% per year for 30 years. In our model, this implies that α_1 falls from 0.475 to 0.235. HLT study the transitional dynamics of this shock. We simply compare the steady-state of the economy before and after the shock has occurred. Table A.3 reports various statistics in both steady-states in our model and in our replication of HLT.

The relative rise in $\frac{R^2}{R^1}$ is the only moment that we use for our calibration. Table A.3 shows that the response of college attendance for all ability types is also similar in our model and in that of HLT. In both models, investment in human capital is somewhat lower in the second steady-state, as shown by the decline in the terminal values of human capital for each type. Finally, Table A.3 shows that there is no change in retirement ages in response to SBTC. In HLT this is an exogenous feature of the model. In our model, this occurs despite the fact that agents are free to choose any retirement age that they wish. The reason that they continue to retire at 65 is that the design of the social security system strongly disincentivises working beyond this age.

B Computational Details

B.1 Solving the Agent's Problem

Conditional on a retirement age, we solve the agent's problem using a "shooting algorithm", exactly as in the original HLT paper:

For a household of age n, type θ, schooling level S, physical capital K, human capital H, who will retire at age $\hat{n}$, the recursive problem is:

$$V_n^{\hat{n}}(H,K,S,\theta) = \max_{C,I,K',H'} U(C) + \delta V_{n+1}^{\hat{n}}(H',K',S,\theta) \tag{39}$$
$$\text{subject to}$$
$$K' = K(1+(1-\tau_k)r) + (1-\tau_l-\tau_P)R^S He(n)(1-I) - C$$
$$H' = H + A^S(\theta) I^{\alpha_S} H^{\beta_S}$$

The FOCs for this problem are:

$$[C]: \quad U'(C_n) = \lambda_n \tag{40}$$
$$[I]: \quad (1-\tau_l-\tau_P)R^S H_n e(n)\lambda_n = \mu_n A^S(\theta)\alpha^S I_n^{\alpha^S-1} H_n^{\beta^S} \tag{41}$$
$$[H']: \quad \delta V_{H,n+1} = \mu_n \tag{42}$$
$$[K']: \quad \delta V_{K,n+1} = \lambda_n \tag{43}$$

where λ_n and μ_n are the Lagrange multipliers associated with the physical capital and human

Table A.3: HLT Experiment: Skill-Biased Technical Change

	Our Model		HLT Model	
	Baseline	SBTC	Baseline	SBTC
r	0.0588	0.0599	0.0588	0.0609
R^1	2	2.27	2	2.20
R^2	2	2.45	2	2.39
Utilized Human Capital $\bar{H}_1$	249	94	274	119
Utilized Human Capital $\bar{H}_2$	287	459	280	446
Physical Capital $\bar{K}$	5605	6849	5725	6814
College Attendance ($\theta = 1$)	0.11	0.47	0.09	0.38
College Attendance ($\theta = 2$)	0.34	0.77	0.28	0.67
College Attendance ($\theta = 3$)	0.56	0.90	0.56	0.89
College Attendance ($\theta = 4$)	0.86	0.99	0.81	0.99
	HS/College	HS/College	HS/College	HS/College
Retirement Age ($\theta = 1$)	65,65	65,65	65,65	65,65
Retirement Age ($\theta = 2$)	65,65	65,65	65,65	65,65
Retirement Age ($\theta = 3$)	65,65	65,65	65,65	65,65
Retirement Age ($\theta = 4$)	65,65	65,65	65,65	65,65
Terminal Human Capital ($\theta = 1$)	9.2/13.0	9.0/12.8	9.4/13.5	9.0/12.9
Terminal Human Capital ($\theta = 2$)	11.8/14.4	11.5/14.1	12.1/14.9	11.5/14.2
Terminal Human Capital ($\theta = 3$)	13.2/15.0	12.8/14.7	13.6/15.5	12.9/14.8
Terminal Human Capital ($\theta = 4$)	12.3/17.6	12.0/17.2	12.6/18.2	12.0/17.4

Notes: These numbers are based on our own replication of the HLT model. Their paper reports values of r, R^1, R^2 in the second steady-state of 0.0609, 2.23 and 2.41, respectively. They report terminal values of human capital in Table 1 for the first steady-state that are very close to the numbers we find. Our estimates of utilized human capital in the original steady-state are close to those seen in their Figure 9. θ denotes the four ability types. $\bar{K}$ measures total capital (held by agents and capitalists).

capital accumulation equations. The envelope conditions are:

$$V_{H,n} = \lambda_n(1-\tau_l-\tau_P)R^S e(n)(1-I_n) + \mu_n[1 + A^S(\theta)I^{\alpha^S}\beta^S H_n^{\beta^S-1}] \tag{44}$$

$$V_{K,n} = \lambda_n(1+(1-\tau_k)r) \tag{45}$$

(relative to HLT, the only difference in FOCs and envelope conditions comes in the derivative of the value function with respect to human capital, where $e(n)$ now shows up)

We can use these conditions to find an Euler equation for consumption:

$$U'(C_n) = \delta(1+(1-\tau_k)r)U'(C_{n+1}) \tag{46}$$

We can also get the following equation defining the evolution of μ_n:

$$\mu_n = \delta[\lambda_{n+1}(1-\tau_l-\tau_P)R^S e(n+1)(1-I_{n+1}) + \mu_{n+1}[1 + A^S(\theta)I_{n+1}^{\alpha^S}\beta^S H_{n+1}^{\beta^S-1}]] \tag{47}$$

Given these conditions, for any guess of terminal values of human and physical capital, the above equations allow us to solve the household's problem backwards to the initial period. We can then iterate until the initial values of human and physical capital are equal to their true values. A detailed description of the algorithm is as follows:

1. Guess physical and human capital levels for period $\bar{n}$.

2. In the final period the household will consume all available resources, therefore:

   $$C_{\bar{n}} = K_{\bar{n}}(1+(1-\tau_k)r) + P$$

   Solve the consumption Euler equation backwards to find $\{C_n\}_{n=1}^{\bar{n}}$. Use the consumption FOC to find $\{\lambda_n\}_{n=1}^{\bar{n}}$.

3. In the final period of working life, optimality implies that $I_{\hat{n}} = 0$ and $\mu_{\hat{n}} = 0$, which implies that $H_{\hat{n}} = H_{\bar{n}}$.

4. Set $i = 1$.

5. Given the values of variables for period $\hat{n} - i + 1$, use the equation for the evolution of μ_n to solve for $\mu_{\hat{n}-i}$.

6. Use the following two equations to solve for $H_{\hat{n}-i}$ and $I_{\hat{n}-i}$:

$$H_{\hat{n}-i+1} = A^S(\theta) I_{\hat{n}-i}^{\alpha_S} H_{\hat{n}-i}^{\beta_S} + H_{\hat{n}-i} \tag{48}$$
$$(1-\tau-\tau_P) R^S H_{\hat{n}-i} e(\hat{n}-i) \lambda_{\hat{n}-i} = \mu_{\hat{n}-i} A^S(\theta) \alpha^S I_{\hat{n}-i}^{\alpha^S-1} H_{\hat{n}-i}^{\beta^S}$$

7. If $i < \hat{n}-1$, set $i = i+1$ and return to step 5. Otherwise, move to step 8.

8. Given the series for $\{C_n\}_{n=1}^{\bar{n}}$, $\{H_n\}_{n=1}^{\bar{n}}$, and $\{I_n\}_{n=1}^{\bar{n}}$, use the physical capital accumulation equation to solve for $\{K_n\}_{n=1}^{\bar{n}}$.

9. If K_1 and H_1 differ from required values, update guesses of $K_{\bar{n}}$ and $H_{\bar{n}}$ and return to step 2.

To solve for the optimal retirement age, we simply deploy the shooting algorithm for every possible retirement age and then select the retirement age that maximizes lifetime utility.

B.2 Solving the General Equilibrium Problem

In this paper, we run a number of experiments that require us to find new equilibrium prices. The three prices are (r, R^1, R^2) which represent the real interest rate and the skill prices of high school and college human capital, respectively. They are pinned down by the derivatives of the production function in (29) - (31). In some experiments, we also increase the labor income tax rate above our baseline of 0.15 and allow the additional revenue to be rebated back to households in the form of a lump-sum transfer P_{LS}. In those experiments, we also need to find the lump-sum transfer that satisfies the government budget constraint in addition to the three prices.

This general equilibrium problem necessitates an outer loop over the shooting algorithm described in the previous section. Because of the discrete nature of some of the decisions in our model (such as retirement age and the decision to work or not), commonly-used nonlinear solvers run into issues converging on a price vector. One way in which we deal with these discontinuities is by using a more robust price-finding algorithm. Our strategy relies on a combination of the Nelder-Mead algorithm from the Julia package `Optim.jl` and a dampened bisection algorithm. This combination is useful because the bisection algorithm works very well if the equilibrium mapping is continuous. The incorporation of Nelder-Mead enables us to be more "robust" to discontinuities, even though it is less efficient.

Given an initial guess for the price (and lump-sum transfer) vector, our algorithm proceeds as follows:

1. Run Nelder-Mead for 100 iterations.
2. Use the outcome from Nelder-Mead as the initial guess to the dampened-bisection algorithm, which runs for 200 iterations. We update the price vector guess according to the following:

$$\begin{aligned} P_{LS}^{new} &= 0.9P_{LS}^{old} + 0.1P_{LS}^{out} \\ r^{new} &= 0.9PLS^{old} + 0.1r^{out} \\ R_1^{new} &= 0.98R_1^{old} + 0.02R_1^{out} \\ R_2^{new} &= 0.98R_2^{old} + 0.02R_2^{out} \end{aligned}$$

   where the variables with superscript *out* represent the prices or transfer that would satisfy the first-order conditions or government budget constraint given the aggregates that result from the shooting algorithm. The extent of dampening on the skill prices (R_1, R_2) is different from that of (r, P_{LS}) because the aggregates in the model respond less continuously to changes in the skill prices.
3. When the dampening algorithm reaches 100 iterations, it is restarted from the best guess price vector so far, and the dampening parameters are halved.
4. If the outcome is still not satisfactory, steps 1 through 3 can be repeated using the outcome of the bisection algorithm as the initial guess for Nelder-Mead.

The algorithm stops if at any point the following residual is less than `1e-8`:

$$\text{residual} = \sum_{i=1}^{4} \left(\frac{P_i^{old} - P_i^{out}}{P_i^{old}} \right)^2$$

where the P_i represent each of the four prices.

B.3 Solving the Laffer Curves

The algorithm in the previous section goes a long way towards solving the general equilibrium problem in the face of the non-convexities in the model. However, it does not go far enough at certain points on our Laffer curves. There are two additional modifications we make in order to overcome these difficulties. One is primarily computational and the other has economic substance. We turn to the computational modification first.

B.3.1 Allowing for continuous retirement ages

Interpolation via cubic splines. One type of difficulty arises simply because workers have only a discrete set of choices over retirement age, i.e., they cannot work for only a fraction of a full year. For the Laffer curves that we analyze in the main part of the paper, we allow agents choose a continuous retirement age by using a spline interpolation method as follows:

1. Interpolate $V(\hat{n})$, the value of retiring at a discrete retirement ages $\hat{n}$, using cubic splines.

2. Find the maximum of $V(\hat{n})$ in the window $[\hat{n}^* - 1, \hat{n}^* + 1]$, where $\hat{n}^*$ is the optimal discrete retirement age. This local maximum is the optimal continuous retirement age.

3. Given the optimal continuous retirement age, we linearly interpolate the agent's paths of human capital, investment, consumption, and capital.

This strategy enables us to successfully compute equilibria at most tax rates on the Laffer curve to a small degree of tolerance.

Analytical method. We have also developed an analytical approach as a way to confirm the accuract of the spline interpolation. In essence, the analytical method adds a second stage to the retirement age optimization problem where we allow workers to re-optimize retirement ages in a continuous manner, holding fixed the human capital choices made in the first stage. We can solve for these continuous retirement ages in closed form as follows.

Let the solution to a worker's annual optimization problem be a last period j of working, a final human capital stock h, and a lifetime consumption profile $\{c_t\}_{t=1}^T$. The worker now gets the opportunity to re-optimize, choosing a continuous time of retirement. The only decision from the annual problem that cannot be changed is the human capital trajectory. However, this is not an issue when interpolating around j, since workers do not undertake any human capital investment at the end of their careers. The re-optimized time of retirement should fall in the interval $(j-1, j+1)$. Let $j+n$ be the interpolated optimal retirement age as a continuous choice variable. We can prove that the solution n must lie in $(-1, 1)$.

The interpolation retains the underlying annual structure of our model. Specifically, for any downward interpolation $n \in (-1, 0]$, the annual discount factor at age j applies to all choices within this window of interpolation, as well as constant age-dependent efficiency units of human capital. In the case of an upward interpolation $n \in [0, 1)$, we apply the annual discount factor and efficiency units of human capital at age $j + 1$.

The change n in retirement age alters a worker's after-tax labor income, which we express as a fraction of the present value of the worker's lifetime consumption in terms of goods in period 1 (i.e., the period that the worker enters the economy),

$$\Delta^I(n) = \begin{cases} \dfrac{\left[\dfrac{1}{1+(1-\tau)r}\right]^{j-1} (1-\tau-\tau_P) n R^S \dfrac{H}{1+e^{s(j-a_p)}}}{\sum_{t=1}^{T} \left[\dfrac{1}{1+(1-\tau)r}\right]^{t-1} c_t} & n \in (-1, 0] \\ \dfrac{\left[\dfrac{1}{1+(1-\tau)r}\right]^{j} (1-\tau-\tau_P) n R^S \dfrac{H}{1+e^{s(j+1-a_p)}}}{\sum_{t=1}^{T} \left[\dfrac{1}{1+(1-\tau)r}\right]^{t-1} c_t} & n \in [0, -1) \end{cases}$$

This calculation is valid under both of our social security schemes.

In response to this change in lifetime labor income, the worker wants to change their entire consumption profile proportionally by the same fraction $\Delta^I(n)$. The change in the lifetime utility of consumption becomes:

$$\begin{aligned} \Delta^C(n) &= \sum_{t=1}^{T} \delta^{t-1} \log\big((1+\Delta^I(n))c_t\big) - \sum_{t=1}^{T} \delta^{t-1} \log(c_t) \\ &= \frac{1-\delta^T}{1-\delta} \log\big(1+\Delta^I(n)\big) \end{aligned}$$

The change in lifetime disutility of work is:

$$\Delta^B(n) = \begin{cases} \delta^{j-1} n B & n \in (-1, 0] \\ \delta^j n B & n \in [0, -1) \end{cases}$$

At the optimal n the marginal disutility of working at a small change in career length should equal the marginal utility of the additional consumption:

$$\frac{d}{dn}\Delta^B(n) = \frac{d}{dn}\Delta^C(n)$$
$$= \frac{1-\delta^T}{1-\delta}\frac{1}{1+\Delta^I(n)}\frac{d}{dn}\Delta^I(n)$$

We can now solve out for the optimal choice of n:

$$n = \begin{cases} \dfrac{1-\delta^T}{(1-\delta)\delta^{j-1}B} - \dfrac{\sum_{t=1}^{T}\left[\dfrac{1}{1+(1-\tau)r}\right]^{t-1} c_t}{\left[\dfrac{1}{1+(1-\tau)r}\right]^{j-1}(1-\tau-\tau_P)R^S\dfrac{H}{1+e^{s(j-a_p)}}} & n \in (-1, 0] \\ \dfrac{1-\delta^T}{(1-\delta)\delta^{j}B} - \dfrac{\sum_{t=1}^{T}\left[\dfrac{1}{1+(1-\tau)r}\right]^{t-1} c_t}{\left[\dfrac{1}{1+(1-\tau)r}\right]^{j}(1-\tau-\tau_P)R^S\dfrac{H}{1+e^{s(j+1-a_p)}}} & n \in [0, -1) \end{cases}$$

When implemented, after calculating the optimal discrete retirement age, we calculate the two values of n that correspond to downward and upward perturbations of the discrete retirement age. We then check which of the three options yields the highest lifetime utility to arrive at the final continuous retirement age, which corresponds to an interior solution where $n \in (-1, 1)$. That interior solution can sometimes be a corner solution to the two equations above, at $n = 0$.

When there is no human capital accumulation (other than the decision of whether or not to attend college), the analytical interpolation method yields two solutions that deliver the same lifetime utility – at the discrete retirement age and either the upward or downward perturbation. Moreover, the solution where $n \neq 0$ is coincides with the optimal continuous age that would be chosen if workers were able to optimize continuously in the first stage. This is because without human capital accumulation, there are no other decisions that the worker is stuck with from the first stage of the optimization. Therefore, this method offers a way to compute exact continuous retirement ages in an annual model.

In sum, the three key features that ensure that the solution of our two-step computation of a continuous career length (first, solving an annual problem, and then using the analytical interpolation) is identical to the solution from a single optimization problem in a continuous

Table B.1: Retirement ages of different high school worker types using spline and analytical interpolation methods

	Discrete	Spline	Analytical
Ability 1	30	29.7507	29.7930
Ability 2	36	35.7252	35.7626
Ability 3	39	38.3935	38.3862
Ability 4	36	36.4830	36.5040

Notes: Evaluated at tax rate 0.61 and the equilibrium prices and Prescott transfer associated with this tax rate in the discrete retirement age solution.

career length are the following: (i) homothetic preferences over consumption, (ii) additively separable disutility of working, and (iii) complete markets. (i) enables us to scale the consumption profile up and down, (ii) enables us to separate out the change in the lifetime disutility of work, and (iii) is what gives us one Lagrange multiplier on a single present-value budget constraint.

We use this analytical method as a way to validate our spline approach. In Table B.1, we illustrate how the two methods perform and compare them to the discrete retirement age solution at tax rate 0.61. We choose this tax rate because it lies in a region where the high school workers are practically not investing in human capital and we do not face any problem of worker types wanting to bifurcate into a short and a long career length (see discussion in next section). In each case, we hold fixed the interest rate, skill prices, and transfer at the equilibrium of the discrete case, and report the optimal retirement ages for high school agents of each ability level. We find that the analytical approach yields a solution that is extremely close to what we get with the spline. This result makes us confident in using the spline as our preferred way to interpolate from discrete to continuous retirement ages.

B.3.2 Convexification

Another challenge arises when a given worker type (ability, schooling pair) is indifferent between a short career length with low human capital accumulation and a long career length with high human capital accumulation. In this situation, there are retirement ages do not respond to prices smoothly, and hence it becomes difficult to find a set of prices that satisfy the general equilibrium conditions. In these cases, we use a convexifcation strategy that splits up the worker type who is indifferent between the two career lengths. It works as follows.

1. For a given guess of prices and the lump-sum transfer, we solve the agents' problem as above using the spline approximation method.

2. For each worker type, we check whether the value function over retirement ages has two local maxima

3. We then identify which worker type is "most indifferent" by choosing the type which has the smallest difference between the two local maxima. If this difference is greater than 0.05%, then we do not convexify any of the agents' retirement choices and complete the algorithm in the usual way (skip the next step).

4. With the agent who is most indifferent, vary the fraction that retire at each of the two ages until the general equilibrium conditions are satisfied.

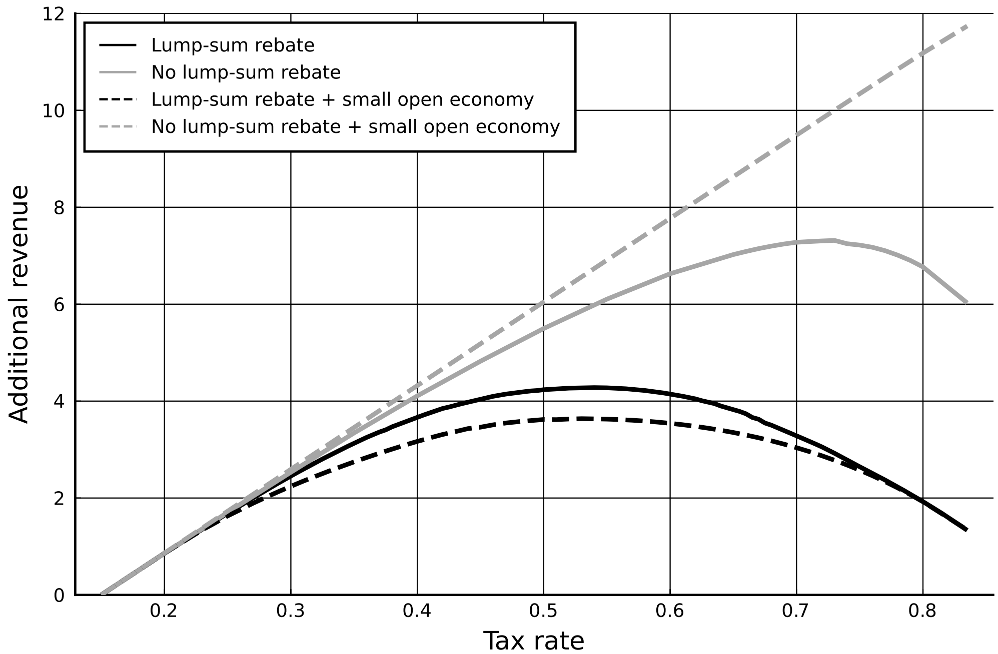
This method introduces smoothness into the general equilibrium problem. This helps us successfully compute equilibrium prices around tax rates where a certain type of agent is nearly indifferent between two much different career lengths. In these situations, as the tax rate rises, the agent type who is indifferent gradually moves from a situation where all of them are choosing the long career length to where all of them are choosing the short career length.

C Taxation Experiments in the Small-Open Economy

In this section, we provide a more comprehensive account of how some of the aggregate outcomes we study in Section 5.1 change in the small-open version of our economy with lump-sum rebates handed back to workers.

Figure C.1 shows the same Laffer curves as in Figure 2(a) from Section 5.1, with the addition of the small-open version of our economy where the tax revenues are given back to households (black dashed line). In the small-open version, where the interest rate is held fixed at its baseline rate, the tax revenues are also hump-shaped but lower, particularly in the middle-range of tax rates. Overall, we find that many of the patterns from the environment with lump-sum rebates carry over to the small-open version, but occur at earlier tax rates. In the small-open economy, the suppression of labor supply occurs earlier, which is why tax

Figure C.1: Laffer curves

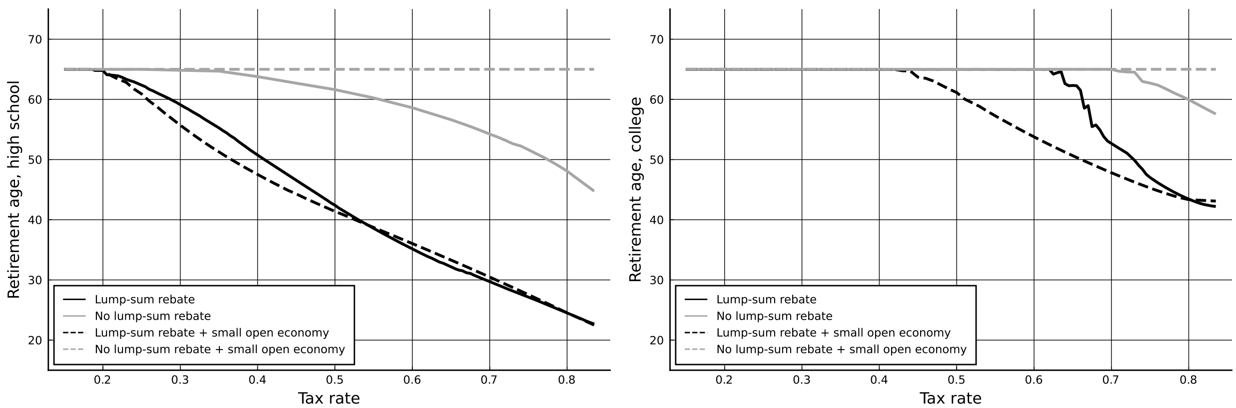


revenues peak at a lower number. The reasons for this will be come clearer from the next set of figures.

Figure C.2 shows the average retirement ages for both high school and college workers. When tax revenues are handed back to households, the small-open economy yields changes that go in the opposite direction to the scenario without lump-sum rebates. High-school and college workers both begin retiring earlier at lower tax rates. This is a cause of the fall in labor supply which leads to the fall in revenue in Figure C.1. From Figure C.2, we can see that below tax rates of 0.5, this is driven primarily by the shortening of career lengths of high school workers, whereas at higher tax rates, it is driven by the same effect among college workers.

These shortened career lengths in this economy are also accompanied by an earlier fall in the college enrollment rates, shown as the black dashed line in Figure C.3(a). In fact, the entire path is similar to the economy where the interest rate is allowed to adjust – with an initial decline, a flattening, and then another decline – but shifted to the left. The dashed line in Figure C.3(b) shows the corresponding pattern in the ratio of college to high school human capital. Like in the closed economy, it rises and then falls, except in the open

Figure C.2: Retirement ages of high school workers (left panel) and college workers (right panel)

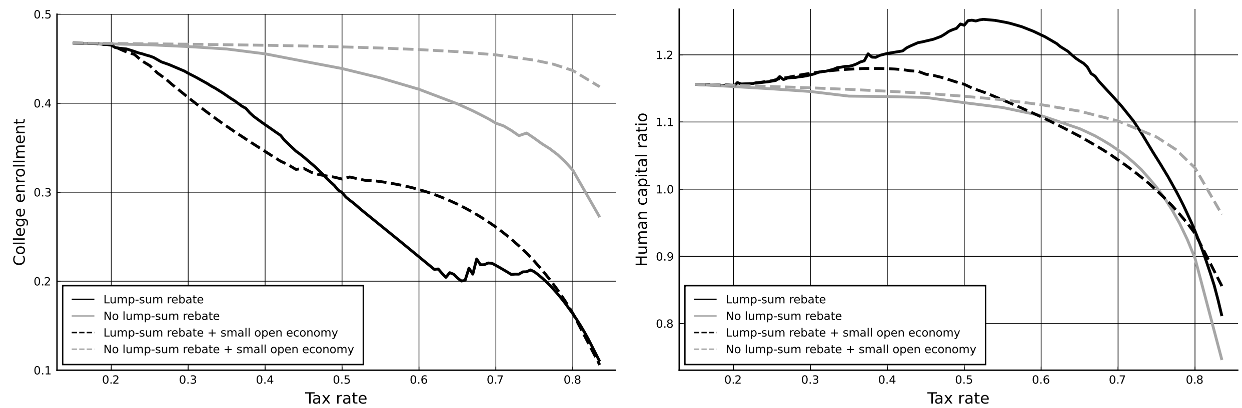


economy the rise is much more muted and the fall occurs earlier. The same is true for the skill premium in Figure C.4: the changes go in the same direction as the closed economy but are muted and start at lower tax rates.

The source of these differences is closely linked to the lack of interest rate response in the small-open economy and the resulting different patterns of capital inflows with and without lump-sum rebates. These patterns are illustrated in Figure 4 in the main text: as discussed there, they look very different from each other. With lump-sum rebates, if the interest rate were allowed to adjust, it would initially decline. This means that agents in the small open economy face a higher interest rate which induces them to start to reduce their labor supply earlier and save more as the tax rate goes up, relative to the closed economy. Initially, the gap between the baseline and general equilibrium interest rate becomes larger, further increasing the disparities between the open and the closed economy. This higher interest rate also reduces the return to going to college relative to the closed economy, which is the force that drives down college enrollment in Figure C.3(a), and dampens the amount of college human capital present as in Figure C.3(b). For the skill premium in Figure C.4, the capital outflows and savings that induce relatively more college than high school human capital decrease the skill premium, but not to the extent of the closed economy in which the adjustment of the interest rate further amplifies these effects.

Things change at high enough tax rates. Like in the closed economy, the transfer becomes large enough to to reduce demand for savings even at the relatively high interest rate. At that point, the outcomes in the small-open economy start to look similar to the closed economy

Figure C.3: College enrollment (left panel) and ratios of college to high school human capital in the production of goods (right panel)



and the two begin to converge.

D Sensitivity Analysis

In this section, we consider alternative calibrations of the function $e(n)$ described in Section A.2. This function determines effectiveness of human capital as agents in the model age. In our baseline calibration we set $\phi_1 = 0.2$ and $\phi_2 = 75$. Figure 1 in the main paper shows that with this calibration, efficiency only begins to decline appreciably as agents reach their 60s. The choice of ϕ_2 implies that efficiency is 50% at the age of 75.

We consider two alternative calibrations, varying both the disutility of labor supply B and the "slope" parameter ϕ_1 simultaneously. We choose the widest range of values of B possible, subject to a constraint that all agents retire at 65 in the original steady-state of the model. In the first, we lower the disutility of labor supply, B, to 0.59 and the slope parameter, ϕ_1, to 0.09. The lower value of ϕ_1 implies that the decline in efficiency is smoother and begins at a significantly younger age, as shown in Figure 1 in the main paper. In the second we raise B to 0.9 and we raise ϕ_1 to 0.31. This implies a sharper decline in efficiency when agents reach their late 60s. Given the requirement that all agents retire at 65 in the original steady-state, these alternatives span the range of possible combinations of B and ϕ_1.

Tables D.1 and D.2 summarize the effect of the HLT skill-biased technical change experiment and our social security reform experiment in our baseline calibration of the model and in these two alternative calibrations. Overall, the alternative calibrations do not significantly

Table D.1: Robustness: HLT Experiment: Skill-Biased Technical Change

	Original Calibration		Low ϕ_1/Low B		High ϕ_1/High B	
	Baseline	SBTC	Baseline	SBTC	Baseline	SBTC
r	0.0588	0.0599	0.0588	0.0602	0.0588	0.0597
R^1	2	2.27	2	2.26	2	2.27
R^2	2	2.45	2	2.45	2	2.45
Utilized Human Capital $\bar{H}_1$	249	94	225	85	253	94
Utilized Human Capital $\bar{H}_2$	287	459	258	413	292	464
Physical Capital $\bar{K}$	5605	6849	5047	9203	5703	6263
College Attendance ($\theta = 1$)	0.11	0.47	0.11	0.47	0.11	0.46
College Attendance ($\theta = 2$)	0.34	0.77	0.34	0.77	0.34	0.76
College Attendance ($\theta = 3$)	0.56	0.90	0.56	0.90	0.56	0.90
College Attendance ($\theta = 4$)	0.86	0.99	0.86	0.99	0.86	0.99
	HS/College	HS/College	HS/College	HS/College	HS/College	HS/College
Retirement Age ($\theta = 1$)	65,65	65,65	65,65	65,65	65,65	62.7,65
Retirement Age ($\theta = 2$)	65,65	65,65	65,65	65,65	65,65	65,65
Retirement Age ($\theta = 3$)	65,65	65,65	65,65	65,65	65,65	65,65
Retirement Age ($\theta = 4$)	65,65	65,65	65,65	65,65	65,65	65,65

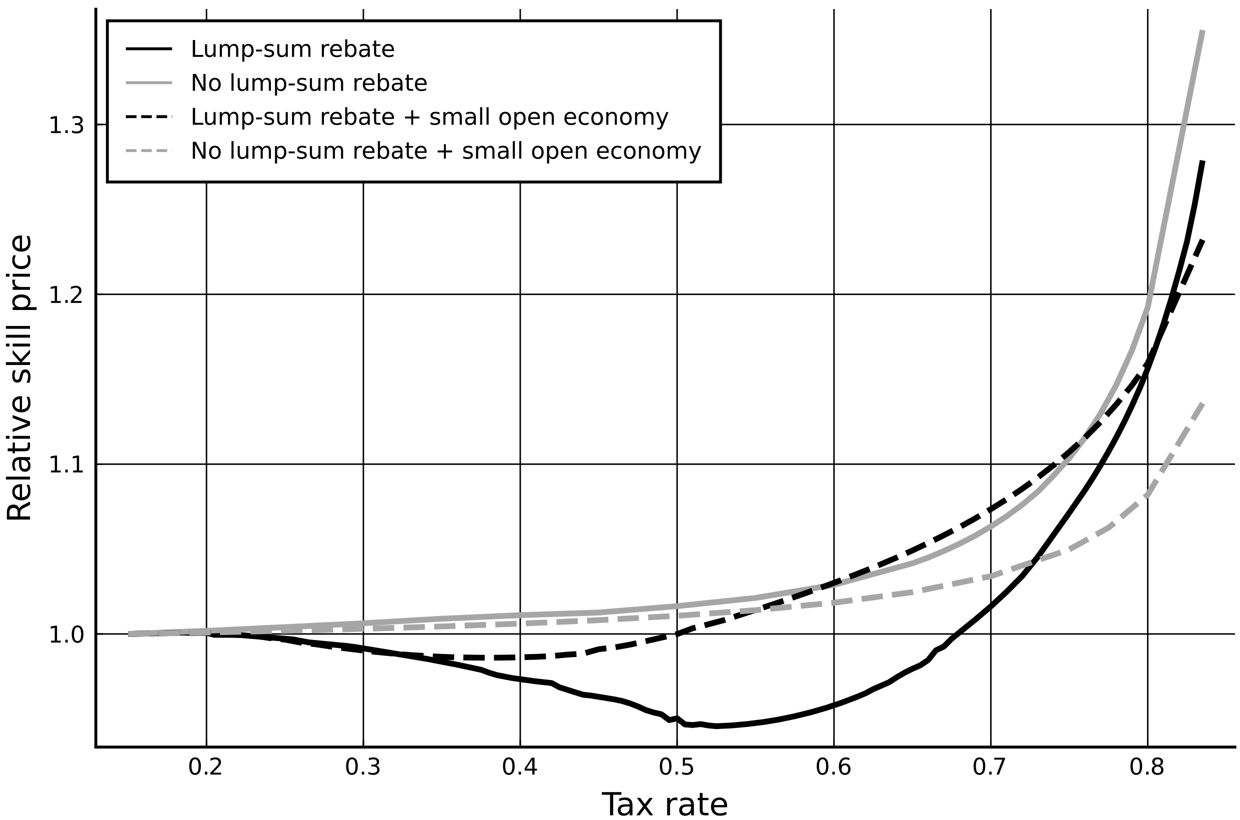
Notes: θ denotes the four ability types. $\bar{K}$ measures total capital (held by agents and capitalists).

Table D.2: Robustness: Social Security Reform

	Original Calibration		Low ϕ_1/Low B		High ϕ_1/High B	
	Baseline	Reform	Baseline	Reform	Baseline	Reform
r	0.0588	0.0617	0.0588	0.0617	0.0588	0.0604
R^1	2	1.97	2	1.97	2	2
R^2	2	1.97	2	1.97	2	1.97
Utilized Human Capital $\bar{H}_1$	249	258	225	237	253	257
Utilized Human Capital $\bar{H}_2$	287	301	258	274	292	305
Physical Capital $\bar{K}$	5605	5583	5047	7514	5703	5072
College Attendance ($\theta = 1$)	0.11	0.08	0.11	0.09	0.11	0.07
College Attendance ($\theta = 2$)	0.34	0.29	0.34	0.30	0.34	0.26
College Attendance ($\theta = 3$)	0.56	0.50	0.56	0.52	0.56	0.47
College Attendance ($\theta = 4$)	0.86	0.83	0.86	0.84	0.86	0.80
	HS/College	HS/College	HS/College	HS/College	HS/College	HS/College
Retirement Age ($\theta = 1$)	65,65	63.3,70.9	65,65	65.3,72.2	65,65	59.9,71.8
Retirement Age ($\theta = 2$)	65,65	64.1,70.9	65,65	65.9,72.1	65,65	61.9,71.8
Retirement Age ($\theta = 3$)	65,65	64.6,70.6	65,65	66.1,72.0	65,65	63.2,71.5
Retirement Age ($\theta = 4$)	65,65	64.5,70.8	65,65	66.1,72.1	65,65	63.5,71.7

Notes: θ denotes the four ability types. $\bar{K}$ measures total capital (held by agents and capitalists).

Figure C.4: Relative price of college to high school human capital



affect the results of these experiments. One area of difference is the implication of social security reform for the retirement age of high school agents. This either shows a slight rise or slight fall, depending on the calibration. In contrast, the retirement age of college agents rises from 65 to 71-72 after the social security reform in all calibrations.

Finally, we consider what would happen in the model if we removed the $e(n)$ function entirely. Figure D.1 shows the retirement ages for each type of agent in such a scenario, for a variety of values of the disutility of labor supply, B. This experiment shows that without human capital depreciation, it is not possible to find a value of B such that both high-school and college educated agents retire at 65 in the baseline version of the model.

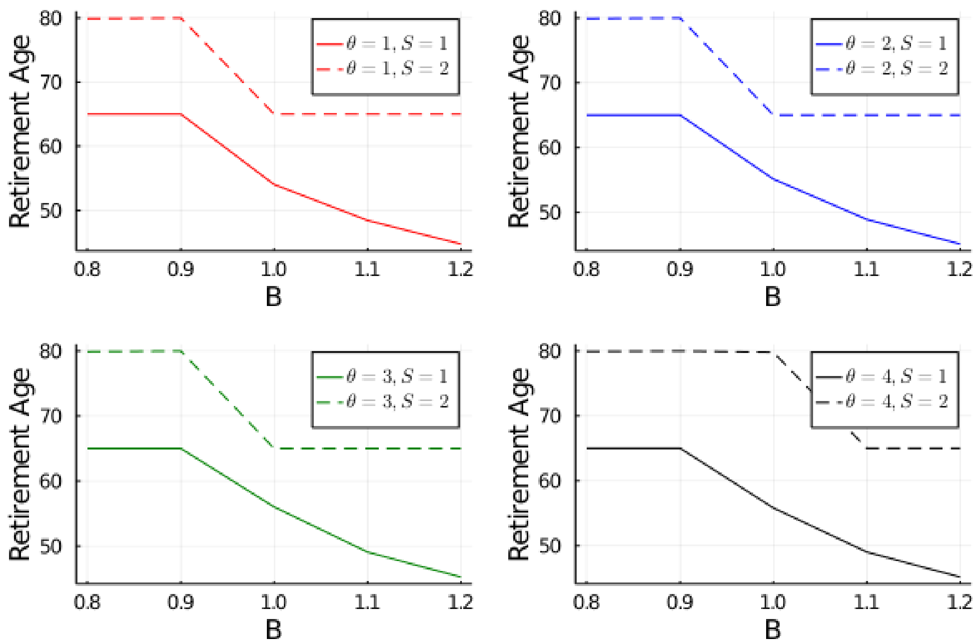
Next, we investigate how responsive labor supply is without human capital depreciation. In Table D.3 we consider the effect of removing the social security system entirely (i.e. setting $\tau_P = 0$ and $P = 0$) in a calibration where B is set to 1. Figure D.1 shows that this value of the disutility of labor supply is the lowest in which the majority of ability types choose

Table D.3: Robustness: Removal of Social Security System

	Original Calibration		No Depreciation	
	Baseline	No Social Security System	Baseline	No Social Security System
B	0.8	0.8	1	1
ϕ_1	0.2	0.2	N/A	N/A
τ_P	0.1	0	0.1	0
P	8	0	8	0
r	0.0588	0.0588	0.0591	0.0591
R^1	2	2.02	2.07	2.07
R^2	2	1.99	1.94	1.94
Utilized Human Capital $\bar{H}_1$	249	289	236	368
Utilized Human Capital $\bar{H}_2$	287	341	300	468
College Attendance ($\theta = 1$)	0.11	0.11	0.04	0.10
College Attendance ($\theta = 2$)	0.34	0.31	0.17	0.29
College Attendance ($\theta = 3$)	0.56	0.51	0.36	0.48
College Attendance ($\theta = 4$)	0.86	0.85	0.72	0.84
	HS/College	HS/College	HS/College	HS/College
Retirement Age ($\theta = 1$)	65,65	68,73	54,65	57,80
Retirement Age ($\theta = 2$)	65,65	69,73	55,65	80,80
Retirement Age ($\theta = 3$)	65,65	69,73	55,65	80,80
Retirement Age ($\theta = 4$)	65,65	69,73	55,80	80,80

Notes: θ denotes the four ability types. In each case, we remove the social security system under a "small open economy" assumption, holding the interest rate constant.

Figure D.1: Retirement Ages Varying Disutility of Labor Supply



an interior solution for their retirement age if they attend college.[28] We do this in a "small open economy" environment, holding the interest rate constant, and compare the results in a model with no depreciation with those that would occur under our original calibration. In the model with human capital depreciation the effect of this experiment is to increase retirement ages by around 4 years for agents with a high school education and 8 years for those with a college education. The effect in the model without human capital depreciation is much more extreme: all agents except the lowest ability high school agent choose to never retire.

[28]The highest-ability college educated agents are the only type that choose to never retire in the "baseline" for this experiment.