Optimal education subsidy and income tax with incomplete markets

Hanson Ho*

July 25, 2020

Abstract. This paper studies the optimal coverage of college education. Considers a social planner who decides who attends college and who pays tax to finance the budget. Sending students to college increases their wages and hence welfare. But increasing college attendance reduces the equilibrium wage and the overall return. Financing the budget through taxation also causes welfare loss. Because of these trade-offs, only students with a sufficiently high return to education should attend college. This paper explores these trade-offs using an overlapping generations model, and determines the optimal coverage of college education and the corresponding tax structure. A simplified model shows that while the current college coverage is close to optimality, the optimal tax policy should be more progressive.

1 Introduction

College is a risky investment. Students pay tuition and then enjoy a wage gain with some probability upon graduation. With complete markets (e.g., credit and contingent markets), such investment should be taken whenever the expected net present value is positive. But due to many market frictions and imperfections, students may not attend college even if the expected net present value is positive. A subsidy (e.g., in a form of reduced tuition or subsidized loan) may overcome this inefficiency and improve welfare. However, such a subsidy is a government expense, which is financed by income tax and may cause distortion. Furthermore, overdoing it may cause an oversupply of graduates and wages to fall. Therefore, the optimal policy should be a balance between the welfare gain of increasing college attendance and the efficiency cost of income tax. Intuitively, due to these trade-offs, only students who benefit sufficiently from education should receive a subsidy. This paper uses an overlaping generations model to analyze these trade-offs and determine the optimal scope of the college subsidy and the corresponding tax policy.

*University of Chicago. Email: hohanson@uchicago.edu

Many studies have discussed the consequences of market imperfections around human capital investment, for instance credit constraint (Lochner and Monge-Naranjo 2011) and the uncertain education outcome (Levhari and Weiss 1974; Altonji 1993). There is also a vast literature on optimal taxation in relation to different departures from certainty and homogeneity such as income shock (Conesa and Krueger 2006) and heterogeneity in labor supply elasticity (Karabarbounis 2016). In this paper, I determine the optimal income tax and school subsidy as a social planner problem, focusing on the property that education augments wages in a heterogeneous and uncertain manner. A simplified version of the model suggests that while the current education policy is close to optimality, the current tax policy is less progressive than the optimal one. Extensions including incorporating a dynamic and stochastic equilibrium framework and an evolving demographic structure are also discussed.

This paper is organized as follows. Section 2 presents a model. Section 3 calibrates the model, discusses the optimal policy, and explores the comparative statics. Section 4 proposes some generalizations of the framework for further research.

2 Model

2.1 Environment

Time is discrete. Every period, a unit mass of new, age-0 agents is born. All age-0 agents survive to age 1. For $n > 1$, only a fraction μ_n of age-n agents survive to age $n+1$. Assume $\mu_n > 0$ for $n = 1, \dots, N-1$ and $\mu_N = 0$.

Every newborn agent i independently faces an investment opportunity: attend college, which costs e and raises productivity by R_i with probability p_i in the future. Let e be an exogenous constant and F be the distribution of (R_i, p_i).

Let w be wage rate. If agent i does not attend college, then the wage income is $wh_i\ell_{in}$ at $n = 0, 1, \dots, N-1$, where ℓ_{in} is labor supply and $h_i = 1$ is productivity. If an agent attends college, then the wage income is zero at $n = 0$, and $wh_i\ell_{in}$ at $n = 1, \dots, N-1$, where $h_i = 1 + R_i$ with probability p_i and $h_i = 1$ with probability $1 - p_i$. Assume N is retirement age such that $\ell_{iN} = 0$.

2.2 Individual value function

Let k be asset and $u(c, \ell)$ be utility where c is consumption and ℓ is labor supply. Let β be subjective discount factor. Let r be interest rate. The value function at age $n \geq 1$ is

$$v_n(k; h, r, w) = \max_{c,\ell} u(c, \ell) + \beta\mu_n v_{n+1}(k'; h, r, w) \tag{1}$$

where $k' = (1+r)k + wh\ell - c$. Note that the terminal value

$$v_N(k; h, r, w) = u(k, 0) \tag{2}$$

can be directly evaluated. By backward induction, we can evaluate $v_n(k; h, r, w)$ for $n = 1, \ldots, N-1$. We also obtain a policy function for asset, $k'^* = \kappa(k; h, r, w)$ and the corresponding optimal consumption $c^*(k; h, r, w)$ and labor supply $\ell^*(k; h, r, w)$. Moreover, given an initial asset, we can use the policy functions to determine the life cycle path of asset, consumption, and labor supply. By construction of the population structure, the sum of these variables over the life cycle is also the cross-sectional aggregate value.

At age $n = 0$, the value of attending college is

$$v_0^c(k; e, R, p, r, w) = \max_c u(c, 0) + \beta(pv_1(k'; 1+R, r, w) + (1-p)v_1(k'; 1, r, w)), \tag{3}$$

where $k' = (1+r)k - e - c$. The value of not attending college is

$$v_0^{nc}(k; r, w) = \max_{c,\ell} u(c, \ell) + \beta v_1(k'; 1, r, w) \tag{4}$$

where $k' = (1+r)k + w\ell - c$.

2.2.1 CES utility

Consider the constant elasticity of substitution utility

$$u(c, n) = \frac{c^{1-\phi}}{1-\phi} - \gamma\frac{n^{1+\psi}}{1+\psi} \tag{5}$$

where ϕ is the coefficient of relative risk aversion, and ψ is the Frisch elasticity of labor supply (inverse).

In Equation (1), given the current asset k and the future asset k', the optimal consumption c and labor supply ℓ can be determined by maximizing

$$\frac{c^{1-\phi}}{1-\phi} - \gamma\frac{n^{1+\psi}}{1+\psi} \tag{6}$$

subject to $k' = (1+r)k + wh\ell - c$. The solution is

$$c = \lambda^{-\frac{1}{\phi}} \quad \text{and} \quad n = (wh\lambda/\gamma)^{\frac{1}{\psi}} \tag{7}$$

where λ satisfies

$$wh(wh\lambda/\gamma)^{\frac{1}{\psi}} - \lambda^{-\frac{1}{\phi}} = k' - (1+r)k. \tag{8}$$

Denote the left-hand side as $\Lambda(\lambda)$. Since Λ is strictly increasing, $\Lambda(0) = -\infty$ and $\Lambda(\infty) = \infty$, there exists a unique positive solution, which can be easily determined numerically.

Now we have determined the optimal consumption c^* and labor supply ℓ^* as a function of current and future assets. Equation (1) becomes

$$v_n(k) = \max_{k'} u(c^*, n^*) + \beta\mu_n v_{n+1}(k'), \tag{9}$$

which gives the policy function for assets. Then the policy function for consumption and labor supply can be determined.

2.3 Equilibrium

Suppose there is a representative firm with production function $AK^\alpha L^{1-\alpha}$ where A is firm productivity, K is total capital, and L is total labor supply. Profit maximization implies

$$\frac{rK}{\alpha} = \frac{wL}{1-\alpha}. \tag{10}$$

Normalize output price to 1. The equilibrium interest rate r and wage rate w clear the labor market and capital market,

$$K = \sum_n \int k_{in}\,\mathrm{d}i \tag{11}$$

$$L = \sum_n \int h_i \ell_{in}\,\mathrm{d}i. \tag{12}$$

This also implies output market clearing.

2.4 Social planner problem

A social planner allocates college places, which can be described by a function Ω where agent i attends college if and only if

$$\Omega(k_i, R_i, p_i) \geq 0. \tag{13}$$

The social planner also sets a labor income tax rate τ and deductible $\underline{w}$, such that agent i pays

$$\tau \max\{0, wh_i\ell_i - \underline{w}\}. \tag{14}$$

Let G be exogenous net public expenses other than tuition cost and labor tax. Let θ be the share of tuition subsidized by the social planner, such that the cost born by students becomes $(1-\theta)e$. The budget constraint is

$$G + \theta e \int \mathbf{1}\left[\Omega(k_i, R_i, p_i) \geq 0\right] \mathrm{d}i = w \sum_n \int h_i \ell_{in} \, \mathrm{d}i. \tag{15}$$

The agents' capital law of motion becomes

$$k' = (1+r)k + wh\ell - \tau \max\{0, wh_i\ell_i - \underline{w}\} - c. \tag{16}$$

Suppose the social planner takes a utilitarian approach and maximizes the total value of all agents, the objective is

$$\begin{aligned}\mathcal{G}(\Omega, \tau, \underline{w}) = \int & v_0^c(k_i; R_i, p_i, r, w, \tau, \underline{w}) \mathbf{1}\left[\Omega(k_i, R_i, p_i) \geq 0\right] \\ & + v_0^{nc}(k_i; r, w, \tau, \underline{w}) \mathbf{1}\left[\Omega(k_i, R_i, p_i) < 0\right] \mathrm{d}i \end{aligned} \tag{17}$$

More generally, we may assign a welfare weight ω_i to each agent i.

2.5 Simplification

For purpose of numerical analysis, we simplify the model and leave the general case to future inquiry. Let $\mu_n = 1$ for $n = 1, \ldots, N-1$. Let r and w be exogenous. Let p_i be fixed such that the only heterogeneity is R_i, and F becomes the distribution of R. Let $\theta = 1$ such that college is fully funded.

The government policy under this setting can be characterized by two thresholds, R_1 and R_2 where $0 \leq R_1 \leq R_2$ ensures people earning unskilled wages are not taxed. Agents with $R \geq R_1$ attend college, and those with $R \geq R_2$ also pay tax if they enjoy a wage gain, such that R_2 proxies tax progressivity.

The total college expense is $(1 - F(R_1))e$. The tax revenue is

$$p \int_{R_2}^{\infty} \tau(1+R) w \ell(k, 1+R, r, (1-\tau)w) \, \mathrm{d}F, \tag{18}$$

The budget constraint becomes

$$p \int_{R_2}^{\infty} \tau(1+R) w \ell(k, 1+R, r, (1-\tau)w) \, \mathrm{d}F = G + (1 - F(R_1))e. \tag{19}$$

The objective becomes

$$\begin{aligned}\mathcal{G}(R_1, R_2, \tau, \underline{w}) = \int_0^{R_1} v_0^{nc}(k; r, w, 0) \, \mathrm{d}F + \int_{R_1}^{R_2} v_0^c(k; e, R, p, r, w, 0) \, \mathrm{d}F \\ + \int_{R_2}^{\infty} v_0^c(k; e, R, p, r, w, \tau) \, \mathrm{d}F. \end{aligned} \tag{20}$$

3 Analysis

In the following, I first determine the model parameters through a combination of internal and external calibration. Using the calibrated parameters, I can return to the social planner problem and determine the optimal policy.

3.1 Parameter estimation

This model neglects childhood and assumes the life cycle of agents begins with one period of college, if they decide to attend. Therefore one period in the model corresponds to 4 years in reality. Suppose people then work for 40 years, this corresponds to 10 periods, hence $N = 10$. Assume agents are economically inactive afterward. Take 30 percent college dropout rate, $p = 0.7$ (NCES 2020b). Normalize $w = 1$. Also set $r = 0.1$ such that the annual interest rate is around 2.5 percent.

For the CES utility

$$u(c, n) = \frac{c^{1-\phi}}{1-\phi} - \gamma \frac{n^{1+\psi}}{1+\psi} \tag{21}$$

we set the coefficient of relative risk aversion to $\phi = 1.5$ (Attanasio and Weber 1995). Meghir and Phillips (2008) surveys the Frisch elasticity of labor supply and points to a value of around 1/3, hence I set $\psi = 3$ to reflect this. The subjective discount factor β, the labor disutility γ, and the initial assets are calibrated.

In our model, R is associated with wage income. Assume $R \sim \text{Lognormal}(\mu, \sigma^2)$, and the parameters are calibrated such that the top 40 percent achieves an average return to education that matches the college-high school wage gap of around 80 percent. Take the tax rate as 20 percent, paid by the top 40 percent of the agents (Elmendorf 2013). That is, $\tau = 0.2$ and $F(R_2) = 0.6$. The net exogenous government spending G and the tuition cost K are calibrated.

In the calibration, I target the following moments: the capital-to-labor income ratio is 2 and the average return to college for those who attend is 80 percent (Autor 2014). Also, let the public sector be 30 percent of the total output, college spending and labor income tax be 10 and 20 percent of the public sector, so they amount to 3 and 6 percent of GDP respectively (OECD 2020; NCES 2020a; CBPP 2020), and G as the difference plus 3 percent budget deficit is 0.06 of the total income. The calibrated values of the model parameters are summarized in Table 1 and the targeted moments in Table 2.

3.2 Policy implications

Now I use the calibrated parameters and solve the government problem, namely to determine the optimal R_1 and R_2 where $0 < R_1 < R_2$, and the corresponding tax rate τ subject to the budget constraint.

For the optimal scope of the subsidy and the tax policy, Figure 1 plots the graph of $\mathcal{G}(R_1, R_2)$ for the calibrated parameters in Tables 1 and 2. The location of the optimal point is marked on the heat map underneath. A direct search yields the optimal point $R_1 = 1.4237$ and $R_2 = 1.6098$, which correspond to the 55th percentile and the 70th percentile. The tax rates that correspond to policy (R_1, R_2) are shown in Figure 2. At the optimal (R_1, R_2), the tax rate is $\tau = 0.4297$.

Compared to the policy parameters used in the calibration, the coverage of college is quite close to the optimal, with around 40 percent college attendance rate. However, the optimal taxation is more progressive than the status quo. Indeed, the tax policy in our model is more rigid than in reality. There are only two tax brackets: $[0, R_2]$ with zero tax and $[R_2, \infty]$ with tax rate τ. But while the optimal policy is roughly a 40 percent tax on the richest 30 percent of people, the current tax policy is approximated by a 20 percent tax on the richest 40 percent of people. Even if we consider the top 1 percent, the effective tax rate of around 30 percent is still below the optimum (Elmendorf 2013). Hence the model supports a more progressive tax structure.

3.3 Optimal policy

Figure 1 shows $\mathcal{G}$ is concave, and both the optimal R_1 and the optimal R_2 are in the interior. Intuitively, when the government expands the scope of the education subsidy by reducing R_1, those who were attending college bear part of the cost by paying a higher income tax and hence are worse off, but those who are marginally admitted to college under the new policy benefit because they now earn a higher income. A similar trade-off exists for expanding the coverage of the income tax. If the government reduces R_2, those who were paying tax enjoy a lower tax rate because more people are sharing the tax liabilities. But this gain is at the expense of those who are now marginally covered by the tax policy.

At the optimal point, expanding the scope of the education subsidy reduces the welfare of those who were already attending college the same as the gain of those who are now marginally admitted to college; and expanding the coverage of the income tax reduces the welfare of those who are marginally taxed the same as the gain of those who were already paying tax.

However, the government objective $\mathcal{G}(R_1, R_2)$ may not attain its maximum in the interior. The solution also depends on the value of the parameters. If the overall return to college is too low, then the government may optimally not subsidize anyone. The restriction $R_1 \leq R_2$ ensures people earning unskilled wages are not taxed. Depending on the labor supply, this constraint may become binding for the government to maintain a budget balance, so it is possible to hit the diagonal and have $R_1 = R_2$.

3.4 Comparative statics

The optimal policy changes with the underlying parameters. Suppose there is an exogenous decline in government spending such that G decreases. Assume assets and prices are not affected. Due to the smaller public expense, the government is more financially capable of funding college education. Therefore the education expense shall be increased, that is, the optimal R_1 decreases. On the other hand, the reduction in public expense relaxes the government budget, which allows the government to contract the tax base. The optimal R_2 increases. Figure 3 supports this, that the peak of $\mathcal{G}$ shifts towards a lower R_1 and higher R_2.

Suppose the mean of the distribution of R is larger. As education becomes more beneficial, it is natural to expect the government optimally reduces R_1 to popularize college education. Due to the larger R, the wage income is higher, so the tax base can be contracted, hence the optimal R_2 should increase. This is confirmed by Figure 4 that the peak of $\mathcal{G}$ shifts towards a lower R_1 and a higher R_2 compared to the benchmark case in Figure 1.

4 Extensions

This paper considers a simple social planner problem in which a utilitarian government balances the trade-off between the welfare gain of college expansion and the efficiency cost of financing the expense by deciding whose tuition to subsidize and whose income to tax. A simplified model shows that although the coverage of college education is about right, a more aggressive tax structure can lead to a welfare gain. Nonetheless, there are several limitations. Going forward, addressing the following issues can enhance the robustness of the results.

4.1 General equilibrium effects

This paper presents a partial equilibrium analysis in the sense that the wage rate and the interest rate are taken as exogenous. Thus changing assets and labor have no general equilibrium effect. This gives us a huge computational advantage. More realistically, for instance, expanding the scope of education subsidy, say by reducing R_1, increases the labor supply and reduces wages. This harms those who were already attending college because now they face more intense competition. But if the government simultaneously reduces R_2, that is, more people are paying tax, then the tax rate can be made lower, which benefits those who were attending college. A full analysis of these two effects and their interactions requires a general equilibrium framework, which is not in the current scope.

Intuitively, with this general equilibrium effect, the negative externality of education is amplified, hence the optimal R_1 would be larger.

4.2 Substitution between skilled and unskilled labor

A related extension is an imperfect substitution between skilled and unskilled labor. In a general equilibrium context, the effect of sending more students to college does not only cause a decrease in wages but also a decrease in skilled wages and a relative increase in unskilled wages. Bringing this effect to the model would probably amplify the negative externality of education and lead to a larger R_1, that is, sending fewer people to college. The quantitative effect depends on the technical substitution between skilled and unskilled labor.

4.3 Dynamic and stochastic income process

In our model, we assume that wage income is constant and deterministic after graduation. This simplifies the calculation. However, the changes in wages over time and the uncertainties involved have important effects on people's decisions to attend college. These also have effects on the optimal policy because uncertainty in income makes it more welfare-improving to raise taxes and redistribute income. We may model productivity as a stochastic process, such that the continuation value term in the individual value function becomes

$$\mathbb{E}\left[v_{n+1}(k'; h', r, w) \mid h\right] \tag{22}$$

Including these elements in our model helps us capture the insurance function of college subsidy and income tax more accurately.

4.4 Capital and consumption taxes

Since I am focusing on human capital investment and the labor market, I consider only labor tax and neglect capital tax. Since capital and consumption taxes also contribute to government revenue, including them in the model can improve the data fit. Further, this allows for a more general tax policy that takes into account the substitution between labor and capital, and also consumption and saving, which are important variables in most life cycle analyses. Suppose there is a linear capital tax τ_k and consumption tax τ_c, the individual budget constraint becomes

$$k' = (1 + (1 - \tau_k)r)k + (1 - \tau)wh\ell - (1 + \tau_c)c. \tag{23}$$

The government budget is modified accordingly.

4.5 College decision as mean-variance optimization

In the numerical analysis, we abstract the variation in the probability of success p. This probability represents the risk of investment in college. Together with the potential wage gain R, this makes the social planner decision Ω akin to the classic portfolio selection problem, with the cost e being the upfront investment cost, and R and p measuring the return and risk of investment. If microdata that enables estimation of these parameters for each college program for each student becomes available, then we may construct the optimal social planner solution more generally.

4.6 Dynamic demographic structure

To simplify the computations, we assume all agents live exactly N periods such that the population structure is constant over time. An advantage of this is that we can compute the cross-sectional aggregate of variables by summing the corresponding variables over life cycle. In the long run, μ_n changes over time. A shifting population structure has important implications for life cycle investment. As many developing countries nowadays are having an aging population, it is important to understand how human capital investment should be adjusted in this environment. The optimal transition dynamics in this environment is also an interesting topic to explore by itself.

References

Altonji, Joseph G. 1993. "The Demand for and Return to Education When Education Outcomes are Uncertain." *Journal of Labor Economics* 11 (1): 48–83.

Attanasio, Orazio P., and Guglielmo Weber. 1995. "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy* 103 (6): 1121–1157.

Autor, David H. 2014. "Skills, Education, and the Rise of Earnings Inequality among the "Other 99 Percent"." *Science* 344 (6186): 843–851.

CBPP. 2020. *Federal Payroll Taxes.* Center on Budget and Policy Priorities.

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

Elmendorf, Douglas W. 2013. *The Distribution of Household Income and Federal Taxes, 2010.* Report 4613. Congressional Budget Office.

Karabarbounis, Marios. 2016. "A Road Map for Efficiently Taxing Heterogeneous Agents." *American Economic Journal: Macroeconomics* 8 (2): 182–214.

Levhari, David, and Yoram Weiss. 1974. "The Effect of Risk on the Investment in Human Capital." *The American Economic Review* 64 (6): 950–963.

Lochner, Lance J., and Alexander Monge-Naranjo. 2011. "The Nature of Credit Constraints and Human Capital." *American Economic Review* 101 (6): 2487–2529.

Meghir, Costas, and David Phillips. 2008. *Labour Supply and Taxes.* Working Paper 3405. London: IZA.

NCES. 2020a. *Postsecondary Institution Expenses.* National Center For Education Statistics.

———. 2020b. *Undergraduate Retention and Graduation Rates.* National Center For Education Statistics.

OECD. 2020. *General government spending.* OECD Data.

A Tables of parameters

Table 1: Model parameters

Parameter	Value	Description
	Externally calibrated	
N	10	Number of periods an agent lives
w	1	Wage rate
r	0.1	Interest rate
ϕ	1.5	Coefficient of relative risk aversion
ψ	3	Inverse Frisch elasticity of labor supply
p	0.7	Probability that college raises income
τ	0.2	Labor income tax rate
$F(R)$	0.6	Fraction of people not attending college
	Internally calibrated	
β	0.9446	Discount factor
γ	0.2332	Labor disutility
a_0	1.0056	Initial asset
G	1.6395	Net government spending
e	1.8258	College cost
μ	-0.9735	Location parameter of distribution of R
σ	0.9133	Scale parameter of distribution of R

Table 2: Targeted moments

Targeted moments	Data	Model
Capital-to-labor income ratio	2.00	2.0122
Labor income tax revenue (% of output)	0.06	0.0559
Net government spending (% of output)	0.06	0.0600
Education expense (% of output)	0.03	0.0300
College wage premium	0.80	0.8276

Table 3: Variable grid

Variable		Grid
Asset	a	$[-5, -4.95, \ldots, 5]$
Return (quantile)	R	$[0, 0.05, \ldots, 0.95]$

Note: A larger grid for asset $[-10, -9.95, \ldots, 10]$ was initially used. The results are the same but the computation time is quadratic in the number of grid points. The grid of return lists the quantiles not the actual values. The actual values of R used in the estimation depends on the distribution parameters (μ, σ). For the values in Table 1, the actual values are $[0, 0.0841, 0.1172, \ldots, 1.2177, 1.6968]$.

B Surface plot of government objective

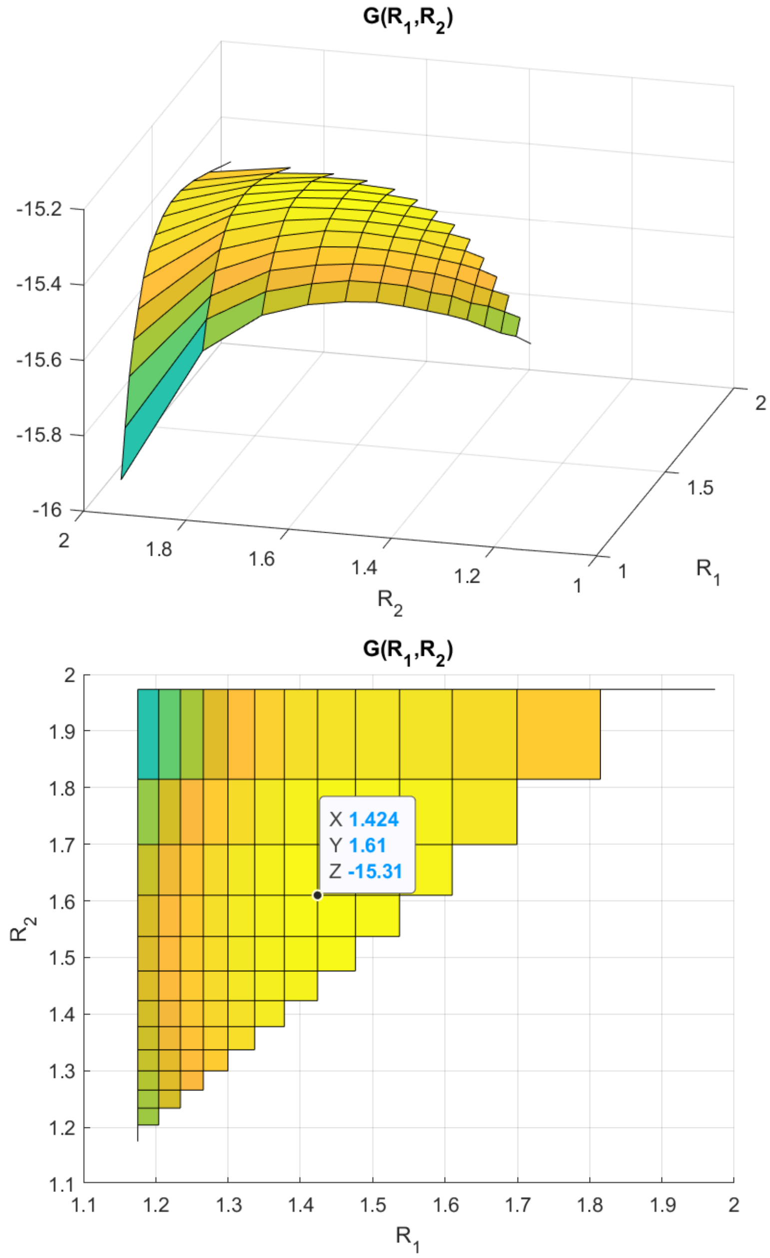

Figure 1: Government objective

Note: Surface plot of $\mathcal{G}(R_1, R_2)$ for parameters in Tables 1 and 2. Objective maximized at $(R_1, R_2) = (1.4237, 1.6098)$ as shown in the heat map.

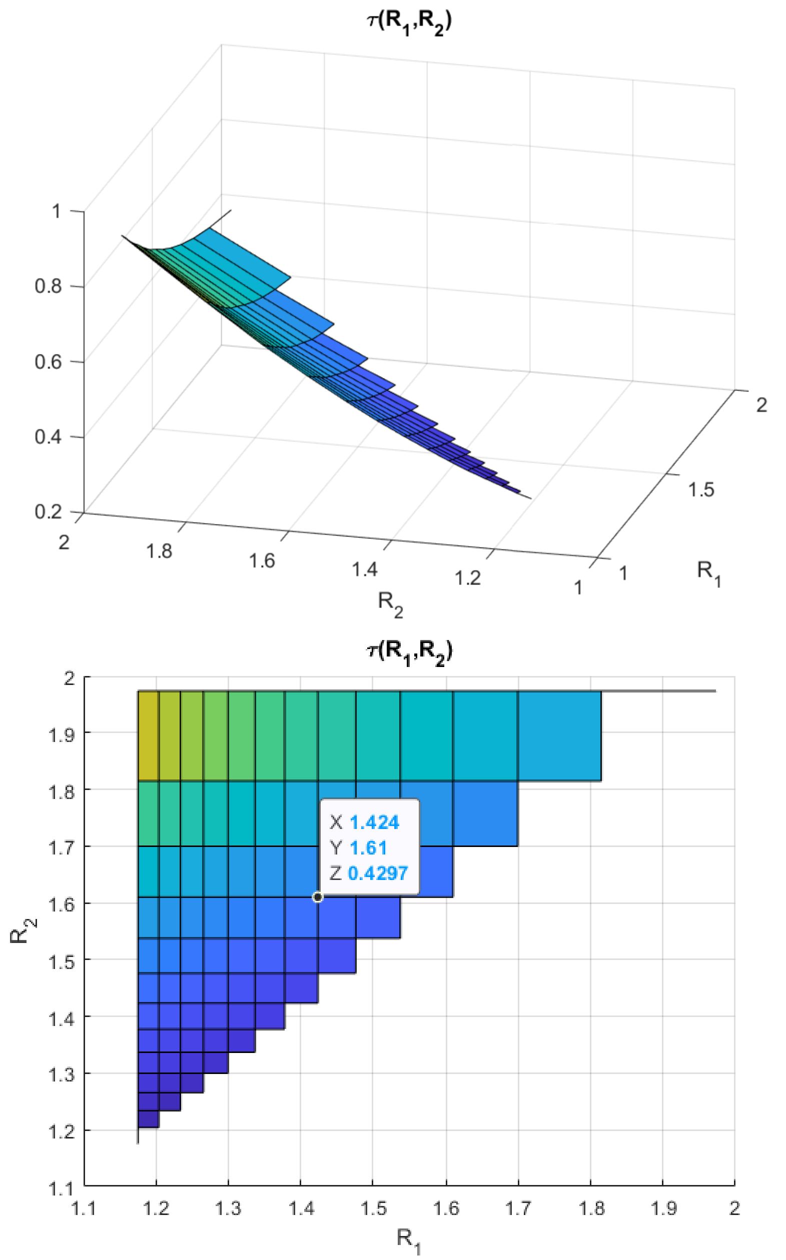

Figure 2: Tax rate

Note: Surface plot of the tax rate that corresponds to every (R_1, R_2) for parameters in Tables 1 and 2. The optimal point $(R_1, R_2) = (1.4237, 1.6098)$ corresponds to $\tau = 0.4297$.

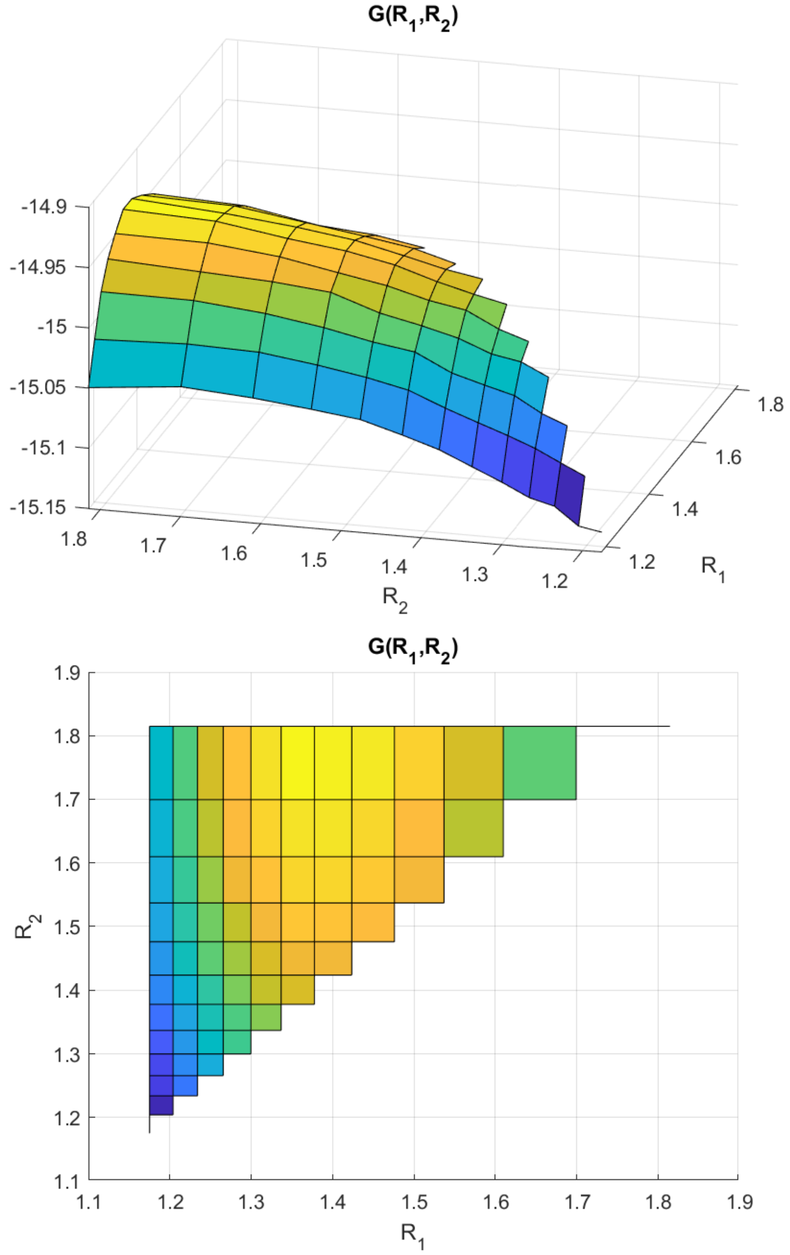

Figure 3: Government objective

Note: Surface plot of $\mathcal{G}(R_1, R_2)$ for parameters in Tables 1 and 2 except G decreases. Optimal R_1 becomes smaller and optimal R_2 becomes larger.

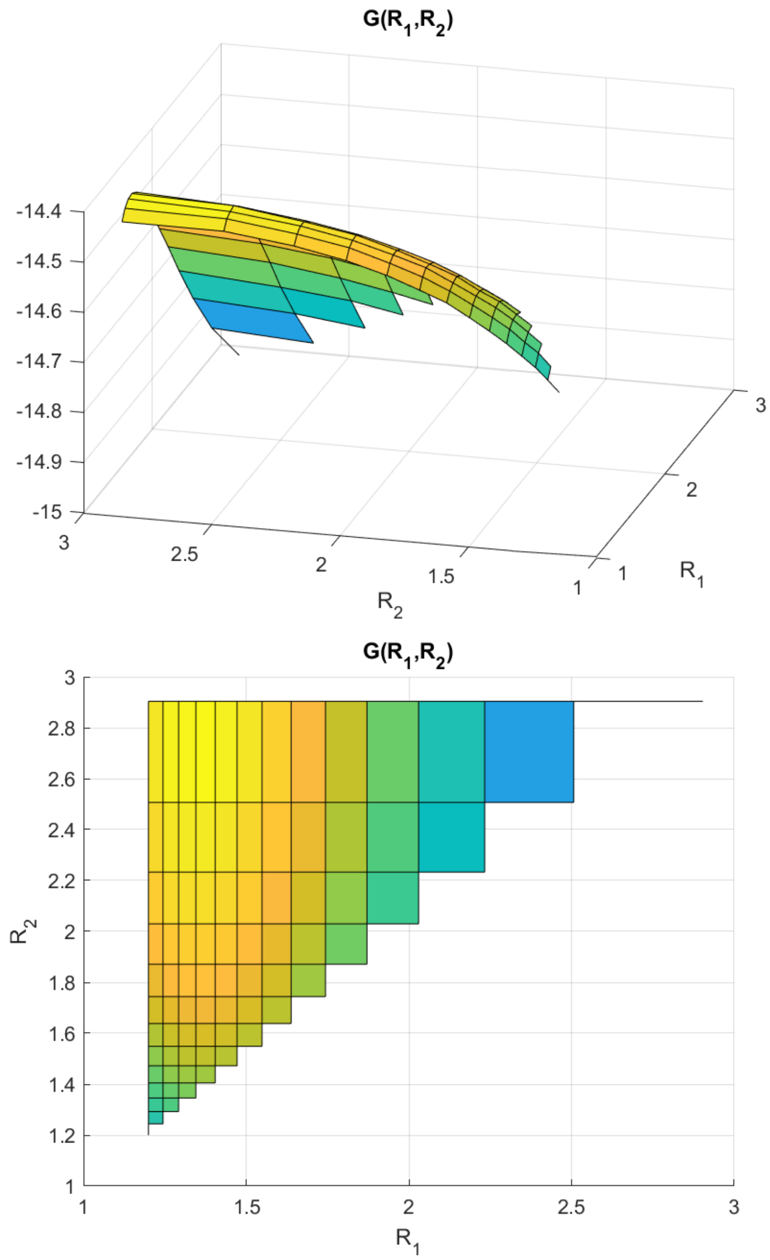

Figure 4: Government objective

Note: Surface plot of $\mathcal{G}(R_1, R_2)$ for parameters in Tables 1 and 2 except μ and σ increase.