Optimal Taxation with Observable Labor Market Entry Shock and Limited Educational Resource

Yizhou Liu

April 22, 2019

Preliminary draft. Please do not circulate.

Latest Version

Abstract

I study optimal taxation and educational policy in an environment in which with limited educational resource and observable labor market entry shocks that persist through life-cycle. Human capital accumulation is unobservable which makes labor decision non-separable. In the optimal policies, the educational resource allocation varies with entry shocks, with low ability agents receiving more while high ability agents receiving less with a good entry shock. The optimal labor wedge decreases over the life-cycle when initial human capital is low to incentivize human capital accumulation. The pattern reverses when initial human capital is high. I calibrate the model to the U.S. data and show that the overall welfare gain is 0.8%. I show how to implement the optimal taxation with an age-dependent labor tax system and an income-contingent student loan repayment design with interest rates indexed to labor market entry shocks. A single income-contingent student loan with interest rate adjustment can achieve 40% of the welfare gain.

Keywords: labor market entry shock, optimal taxation, student loan design.

1 Introduction

It is well documented that labor market entry conditions for college graduates have large and persistent impact on wage: it is estimated that a 5% rise in unemployment implies an initial loss of earnings of about 10 percent which fades to 0 by 10 years. The total loss of accumulated earnings amounts up to 5% (Oreopoulos et al. (2012), Liu et al. (2016), Kahn (2010), Altonji et al. (2016)). This means that the return to educational investment depends crucially on aggregate economic conditions upon labor market entry. How should limited education resource be distributed for cohorts with different labor market entry shocks? How does the educational policy interact with labor taxation when there is unobservable human capital accumulation after graduation? This paper tries to answer these questions.

I study the optimal Mirrleesean taxation in an environment with observable labor market entry shocks and limited educational resource. The model features a Ben-Porath economy with unobservable human capital accumulation after graduation. Agents differ by ability which is fixed but unobservable to the social planner. Using a first-order approach, I derive the optimal labor wedge and educational fee formula. Using a solvable example, I show that the optimal taxation and educational fees are higher for cohorts with good entry shocks.

I calibrate the model to the U.S. economy and show that the educational fee should be 9% higher for cohorts with a 5% labor market entry shock. The increase in educational fee is important: if educational fee is equalized across cohorts, due to limited educational resource, the social planner needs to uniformly decrease the educational resources across agents of different ability, which increases inequality in educational resource allocation.

In the optimal policy, the social planner redistributes the limited educational resource in a boom such that low ability agents experience an increase in educational resource while the opposite holds for high ability agents. This is because the marginal reduction in labor disutility due to education is more sensitive to aggregate shock when education level is low, as is the case for low ability workers. Thus, a 5% entry shock induces reduction in marginal labor disutility that more than offsets the 9% increase in educational fee, causing the low ability workers to choose higher levels of education to equalize marginal benefit and marginal cost. For the high ability workers, the marginal reduction in labor disutility is less sensitive to aggregate shock due their higher level of education attainment. A 9% increase in educational fee hence lead them to reduce their educational choice.

The average optimal labor wedge is decreasing in early periods for low ability agents while increasing for high ability agents. The reason is that the social planner would like to encourage human capital accumulation when initial human capital is low, as is the case for low ability agents. As labor depresses human capital accumulation, tax rates are high in early periods and declines in later periods. The optimal taxation also differs across cohorts in which the cohort with good entry shock faces higher tax rates. For high ability workers, their initial human capital is high, so that human capital accumulation is less important. The tax rates hence increase over the life-cycle due to rising income, resembling the baseline case. Moreover, the difference in tax rates across cohorts is minimal, suggesting that high ability agents is less affected by labor market entry shocks, echoing the finding in Oreopoulos et al. (2012).

Another feature of the optimal taxation is the educational tax deduction, which is decreasing in education attainment. That is, lower education attainment implies higher life-cycle tax deduction, so that the agents with higher educational choice subsidizes the agents with lower educational choice. Educational tax deduction increases in time, suggesting that the benefit of initial human capital accumulates over the life-cycle.

I show how the optimal policies can be implemented by an age-depend tax system and an income contingent student loan repayment design with interests rate indexed to labor market entry shocks. Compared to the baseline, the total income increases by 0.8% under the optimal taxation. I show also that a single income-contingent student loan plus an interest rate indexed to labor market entry conditions can achieve 40% of the welfare gain: income-contingent student loan changes the effective tax rate which brings the allocations close to the optimal. If the interest rate remains constant, agents would like to invest more in a boom, leading to shortage of educational resources. A uniform reduction in educational choice would lead to higher wage inequality and lower total income.

The result emphasizes the need to take into consideration labor market entry shocks in student loan design. Specifically, the current student loan interest rates and repayment rates are determined at the time of loan receipt and do not correlate with aggregate labor market conditions. Thus, cohorts that enter the labor market in a recession are likely to face more difficulties repaying their loans, have higher rate of default, and result in worse labor market outcomes. I also show that in the single income-contingent student loan design, it is likely that low ability agents pay less than the total loan amount while high ability agents pay more than the total loan amount.

The paper is closely related to Kapicka (2015), in which he studies optimal taxation with unobservable human capital accumulation. I augment the model in Kapicka (2015) with limited education resource and heterogeneous cohorts that differ in labor market entry shocks to study the interaction between educational resource allocation, labor market entry shocks, and optimal taxation. In the implementation of the optimal mechanism, I emphasizes the role of income-contingent student loan design and find welfare gains similar to Findeisen and Sachs (2016). I introduce a novel feature of student loan interests rate that is indexed to labor market entry conditions in the implementation. This is important for optimal allocation under limited educational resources.

The rest of the paper is organized as follows: section 2 describes the model. section 3 introduces the planner's problem and solve the optimal taxation and educational policy. section 4 illustrates the optimal policies in a simple three periods, two-cohorts example. In section 5, the model is calibrated to the U.S. data to quantitatively evaluate the welfare improvements. section 6 discusses implementation and section 7 concludes.

2 Model

The model is similar to Kapicka (2015). I augment the model with differential initial human capital based on educational choice and observable labor market entry shocks that persist through life-cycle. There is a continuum of heterogeneous agents with unobservable but fixed ability $\theta \in \left[\underline{\theta}, \overline{\theta}\right] \equiv \Theta$ with pdf $f(\theta)$. There are multiple cohorts differ by labor market entry shocks σ. Labor supply is z_t and output is given by $y_t = \sigma\theta z_t$.

In period 0, agents make educational choice a which together with their ability θ determines their human capital in period 1: $h_1 = h_1(\theta, a)$. Individuals live for T periods. The disutility from labor supply is given by a function $W(z_1, z_2, ..., z_T, h_1)$ where I make explicit the dependence on initial human capital h_1. The disutility function is non-separable in labor supply, which is the case in the Ben-Porath model with unobserved human capital accumulation. Individuals are risk neutral with discount rate β. They maximize life time discounted consumption minus disutility from labor supply:

$$\max \sum_{t=0}^{T} \beta^t c_t - W(z_1, z_2, ..., z_T, h_1) \tag{1}$$

The planner's problem is to choose a vector of allocations $(u(\theta,\sigma), z, a)$ to minimize the cost of delivering lifetime discounted utility $u(\theta, \sigma)$ to agent type θ and cohort σ such that the following conditions are satisfied:

$$\begin{gathered}\text{Promise-Keeping Constraint}: \\ u(\theta,\sigma) \geqslant \underline{U}(\theta,\sigma), \quad \forall\theta\,\forall\sigma\end{gathered} \tag{2}$$

$$\begin{aligned}\text{Incentive-Compatible Constraint}: & \\ u(\theta,\sigma) \geqslant u(\hat{\theta},\sigma) &+ W(z_1(\hat{\theta}), ..., z_T(\hat{\theta}), h_1(\hat{\theta}, a(\hat{\theta}))) \\ &- W(\frac{y_1(\hat{\theta})}{\sigma\theta}, ..., \frac{y_T(\hat{\theta})}{\sigma\theta}, h_1(\theta, a(\hat{\theta})))\end{aligned} \tag{3}$$

Educational Resource Constraint :

$$\int_{\underline{\theta}}^{\overline{\theta}} a(\theta,\sigma) f(\theta) d\theta \leqslant A, \quad \forall\sigma \tag{4}$$

The promise keeping constraint equation (2) can also be thought of as a participation constraint with $\underline{U}(\theta,\sigma)$ given by the agents' outside option. The incentive-compatible constraint differs from Kapicka (2015) in that reporting strategy now also affects education attainment and hence initial human capital h_1. The incentive constraint also differs by cohort, so that for a given type of agent θ, the constraint might bind in one cohort while not bind in the other cohorts. Limited educational resource constraint equation (4) is crucial for the interaction between taxation and educational policy. I assume the total educational resource A is fixed across cohorts.

3 Optimal Policy

I first use a first-order approach to re-write the incentive-compatible constraint as:

$$\begin{aligned} u(\theta,\sigma) = & \int_{\underline{\theta}}^{\theta} \sum_{t=1}^{T} W_{z_t}(z_1,...,z_T,h_1)\frac{z_t}{\epsilon} d\epsilon \\ & + \int_{\underline{\theta}}^{\theta} W_{h_1}(z_1,...,z_T,h_1)\frac{\partial h_1}{\partial \theta}(\epsilon, a(\epsilon)) d\epsilon + u(\underline{\theta},\sigma) \end{aligned} \tag{5}$$

The relaxed incentive compatible constraint has three parts. The first part suggests that the labor supply allocation needs to make the agents reporting truthfully their ability. The second part decrees that educational resource allocation needs to align with individuals' incentives, taking into consideration the effect of initial human capital and labor market entry shock on lifetime income. Under monotonicity conditions on $y(\theta)/\theta$ equation (5) is equivalent to equation (3).

The planner's relaxed problem is to maximize total income subtract by labor disutility, subject to equations (2), (4) and (5):

$$
\begin{aligned}
K(\{U\}_{(\Theta,\Sigma)}) = \min_{\mu,\lambda} \max_{\{z_t\},a} \sum_{i=1}^{N} \int \Big\{ \sum_{t=1}^{T} \Big[& \beta^t \sigma_i \theta z_t - \mu(\theta,\sigma_i) \int_{\underline{\theta}}^{\theta} W_{z_t}(z_1, ..., z_T, h_1) \frac{z_t}{\epsilon} d\epsilon \Big] \\
& - \mu(\theta,\sigma_i) \int_{\underline{\theta}}^{\theta} W_{h_1}(z_1, ..., z_T, h_1) \frac{\partial h_1}{\partial \theta} d\epsilon \\
& - W(z_1, ..., z_T, h_1) + \lambda(\sigma_i)(A - a(\theta,\sigma_i) \Big\} f(\theta) d\theta \\
& + \sum_{i=1}^{N} \int \Big\{ (\mu(\theta,\sigma_i) - 1) u(\theta,\sigma_i) - \mu(\theta,\sigma_i) u(\underline{\theta},\sigma_i) \Big\} f(\theta) d\theta
\end{aligned}
\tag{6}
$$

where $\mu(\theta, \sigma_i)$ is the Lagrangian multiplier on equation (5) and $\lambda(\sigma_i)$ is the Lagrangian multiplier on equation (4).

Changing the order of integration, we obtain the following equivalent formulation:

$$
\begin{aligned}
\mathcal{L} = \min_{\mu,\lambda} \max_{\{z_t\},a} \sum_{i=1}^{N} \int \Big\{ \sum_{t=1}^{T} \Big[& \beta^t \sigma_i \theta z_t - W_{z_t}(z_1, ..., z_T, h_1) z_t X_{\mu}(\theta,\sigma_i) \Big] \\
& - W_{h_1}(z_1, ..., z_T, h_1) \frac{\partial h_1}{\partial \theta} \theta X_{\mu}(\theta,\sigma_i) \\
& - W(z_1, ..., z_T, h_1) + \lambda(\sigma_i)(A - a(\theta,\sigma_i) \Big\} f(\theta) d\theta \\
& + \sum_{i=1}^{N} \int \Big\{ (\mu(\theta,\sigma_i) - 1) u(\theta,\sigma_i) - \mu(\theta,\sigma_i) u(\underline{\theta},\sigma_i) \Big\} f(\theta) d\theta
\end{aligned}
\tag{7}
$$

where

$$
X_{\mu}(\theta,\sigma_i) = \frac{1 - F(\theta)}{\theta f(\theta)} \frac{\int_{\theta}^{\overline{\theta}} \mu(\epsilon,\sigma_i) f(\epsilon) d\epsilon}{1 - F(\theta)} \tag{8}
$$

is the cumulative distortion as in Kapicka (2015) and $F(\theta)$ is the CDF of the ability distribution.

Taking the first order conditions, denote the derivative of labor disutility as W_x and second-order derivative as W_{xy}. I derive the following lemma.

Lemma 1. *1. The labor wedge is given by the following equation:*

$$\frac{\tau_t}{1-\tau_t} = (1+\rho_t+\phi_t)X_\mu \tag{9}$$

where $\rho_t = \sum_{j=1}^{T} \rho_{t,j}$*,* $\rho_{t,j} = \frac{W_{z_t z_j} z_j}{W_{z_t}}$*, and* $\phi_t = \frac{W_{h_1 z_t}\theta\,\partial h_1/\partial\theta}{W_{z_t}}$*.*

2. The educational fee is given by:

$$\begin{aligned}\lambda(\sigma_i) = &-\sum_{t=1}^{T} \frac{\phi_t W_{z_t} z_t X_\mu\, \partial h_1/\partial a}{\theta\, \partial h_1/\partial\theta} \\ &- \left[W_{h_1 h_1}\frac{\partial h_1}{\partial\theta}\frac{\partial h_1}{\partial a} + W_{h_1}\frac{\partial^2 h_1}{\partial\theta\partial a}\right]\theta X_\mu \\ &- W_{h_1}\frac{\partial h_1}{\partial a}\end{aligned} \tag{10}$$

3. The Lagrangian multiplier $\mu(\theta,\sigma)$ *and the promise-keeping constraint have the following relation:*

$$[\mu(\theta,\sigma_i)-1][u(\theta,\sigma_i)-\underline{U}(\theta,\sigma_i)] = 0 \tag{11}$$

The labor wedge expression is similar to Kapicka (2015), with the augmentation of ϕ_t which determines how labor wedge should respond to the relation between initial human capital h_1 and future labor supply z_t. $\partial h_1/\partial\theta$ is the marginal gain in initial human capital due to higher ability. Its effect on labor wedge depends on the sign of the marginal reduction in disutility due to higher initial human capital $W_{h_1 z_t}$. If $W_{h_1 z_t} > 0$, the labor wedge is increasing in the marginal gain of initial human capital in ability.

To understand this result, it is important to first understand the interpretation of $W_{h_1 z_t}$. W_{h_1} is the life-cycle reduction in labor disutility due to higher initial human capital. A positive $W_{h_1 z_t}$ implies that higher labor supply damps the marginal reduction in labor disutility. If initial human capital is increasing in ability θ, high ability agents benefit increasingly less from higher labor supply, making it more difficult more the planner to provide incentives. To counteract that effect, the planner increases labor wedge to reduce labor supply of high ability agents.

On the other hand, a negative $W_{h_1 z_t}$ implies a smaller labor wedge because the planner would like to encourage labor supply because it decreases overall disutility via W_{h_1}. Without initial human capital accumulation from education, this effect is absence and equation (9) reduces to $\tau_t/(1-\tau_t) = (1+\rho_t)X_\mu$, the same as in Kapicka (2015).

ρ_t includes cross derivative between z_t and labor supply in all periods. When labor disutility is separable, all cross derivatives would disappear. With unobservable human capital accumulation, labor disutility is non-separable.

The determination of educational fee consists from three parts. The first part is the cumulative effect of disutility reduction for all future labor supply decisions due to initial human capital. This can be seen clearly if we rewrite the first term on the right-hand side of equation (10) as :

$$-\sum_{t=1}^{T} W_{z_t h_1} z_t X_\mu \, \partial h_1/\partial a \tag{12}$$

The marginal reduction in labor disutility, or marginal benefit of initial human capital in reducing future labor disutility, is given by $-\sum_{t=1}^{T} W_{z_t h_1} z_t$. This term captures the gain of higher initial human capital on labor supply. To obtain the effect of educational attainment, we multiply it by $\partial h_1/\partial a$.

The second term on the right-hand side of equation (10) represents the distributional effect of educational benefit. It is the sum of two parts. The first part captures the gain in educational attainment due to the curvature of the labor disutility function. If the marginal benefit of initial human capital is increasing, so that higher initial human capital implies larger marginal disutility reduction, it would push up the educational fee as agents desire more educational resource. The second part captures the complementarity between ability and education in determining initial human capital. A higher degree of complementarity means higher ability agents receive more marginal benefit from education.

The third term is the marginal reduction in labor disutility due to higher level of education. Because there is no incentive distortion for it, it is not augmented by the cumulative distortion.

Notice that the Lagrangian multipliers $\mu(\theta, \sigma)$ enters the cumulative distortion equation (8). It establishes a link between the promise-keeping constraint equation (2) and the cumulative distortion X_μ. The cumulative distortion is increasing in $\mu(\epsilon, \sigma)$, which implies that the larger the promised utility for agent types $[\theta, \overline{\theta}]$, the more taxes needs to be collected from agents of type θ. The more binding the promise-keeping constraint is, the smaller the Lagrangian multiplier μ hence X_μ. On the other hand, if the promise-keeping constraints never bind, which would be the case for a Rawlsian social planner, the link between the constraint and the cumulative distortion vanishes and X_μ is simply $(1 - F(\theta))/\theta f(\theta)$.

There is also an indirect link between the promise-keeping constraint and the educational fee. Tighter promise-keeping constraint would be a higher educational fee because the cumulative distortion enters equation (10). The intuition is that larger promised utilities require higher level of education, which increases the educational fee because of limited educational resources.

4 An Illustrative Example

To better understand the results, I work out a fully solved example with three periods and two cohorts. I first motivate the non-separable labor disutility function W using the Ben-Porath model as in Kapicka (2015). There is unobservable human accumulation $\{h_{t+1}\}_{t \leqslant T}$ and individuals minimize the cost of labor supply:

$$W(z_1, ..., z_T, h_1) = \min_{\{h_{t+1}\}} \sum_{t=1}^{T} \beta^t V(s_t, l_t) \tag{13}$$

where s_t is human capital accumulation effort and l_t is effective labor effort. $V(s, l)$ is per period disutility of human capital accumulation and effective labor effort. It has strictly positive first order derivatives: $\partial V/\partial s > 0$, $\partial V/\partial l > 0$, $\forall s, l > 0$.

I use the following functional forms: $V = s_t l_t$, $s_t = \exp(h_{t+1} - h_t)$, and $l_t = z_t \exp(-h_t)$. This choice is for the purpose of obtaining analytic solutions. In general, I need to require that human capital accumulation effort is increasing in human capital stock next period h_{t+1} and the effective labor effort is decreasing in human capital stock. With $T = 2$, I derive the labor disutility function as:

$$W(z_1, z_2, h_1) = \frac{1}{2} e^{-h_1} z_1^2 z_2^2 \tag{14}$$

This simple example illustrates how unobservable human capital accumulation leads to non-separable labor supply disutility function.

I further assume that there are two cohorts indexed by $\sigma_h > \sigma_l$. Referring to lemma 1, setting the functional form $h_1 = \theta + a$, the labor wedges reduce to:

$$\tau_1(\theta, \sigma_i) = \tau_2(\theta, \sigma_i) = \frac{(4-\theta) X_\mu(\theta, \sigma_i)}{(4-\theta) X_\mu(\theta, \sigma_i) + 1}, \quad i = h, l \tag{15}$$

The take-away are two-fold. First, due to non-separability, period 1 and period 2 labor supply play equal role in labor disutility function equation (14) even though the agent discount the disutility from effective labor effort in the second period, hence the optimal labor wedge is the same in both periods. If the labor disutility is separable, the agent would discount future labor disutility more so the optimal labor wedge would increase. The second take-away is that labor wedge is increasing in the cumulative distortion $X_\mu(\theta, \sigma_i)$ which in turn depends on the baseline promised utility $\underline{U}(\theta, \sigma_i)$. An egalitarian social planner would equalize the baseline promised utilities across cohorts, making the promise-keeping constraints less likely to bind in a boom, i.e. when σ is larger. The result is an increase in the cumulative distortion X_μ hence higher taxes for cohorts with a good labor market entry shock. I state this in the lemma 2

Lemma 2. *The optimal labor wedge is increasing in labor market entry shock in the sense that cohorts with good shocks face higher labor wedge.*

Educational fee is given by:

$$\lambda(\sigma_i) = \left[(2 - \frac{\theta}{2})X_\mu(\theta, \sigma_i) + \frac{1}{2}\right] e^{-(\theta+a)} z_1^2 z_2^2 \tag{16}$$

In this specific example, the "marginal benefit" of education, which includes both income gain from education and incentive provision, is decreasing in ability and education. To equalize marginal cost and "marginal benefit", the social planner needs to allocate more educational resource to low ability agents. This is easily seen from equation (16) since the right-hand side is decreasing in θ and a.

The educational fee $\lambda(\sigma_i)$ depends on cumulative distortion and labor supply. I show previously that the cumulative distortion is increasing in σ_i. The intuition is that it is easier to satisfy the promise keeping constraint when agents face a good entry shock. It is also the case that labor supply is higher when agents face good shocks, so that the educational fee is also increasing in labor market entry shocks:

Lemma 3. *The optimal educational fee is higher for cohorts with good entry shocks and lower for cohorts with bad entry shocks.*

What if the educational fee is the same for all cohorts? Cohorts with good entry shock would then demand more education. Due to limited educational resource, the social planner has to decrease their educational attainment. I discuss the implication of it in more detail in section 5.

5 Quantitave Analysis

In this subsection, I numerically solve the model and discuss key features on the optimal taxation and how it can be implemented. Similar to Kapicka (2015), I choose the disutility from accumulation unobservable human capital as:

$$V(l,s)=\frac{(l+s)^{1+\nu^{-1}}}{1+\nu^{-1}}$$

Human capital evolves according to:

$$h_{t+1}=G(h_t,s_t)=(1-\delta)h_t+\delta(h_ts_t)^{\alpha}$$

Effective labor effort is $l_t=z_t/h_t$.

The initial human capital depends on educational choice and ability via:

$$h_1=h_1(\theta,a)=\gamma a\log(\theta)$$

The cross-partial derivative is $\gamma/\theta>0$. This means that high ability agents gain more initial human capital through education.

Labor disutility is implicit determined in the Ben-Porath economy equation (13). Due to unobservable human capital accumulation, labor disutility is non-separable in labor supply $\{z_t\}$. This allows non-trivial cross-Frisch elasticities of labor and dependence of all future marginal labor disutility on initial human capital.

In the baseline calibration, I use the following tax schedule given in Kapicka (2015):

$$\tau(y)=\kappa_0\Big[y-\big(y^{-\kappa_1}+\kappa_2\big)^{-1/\kappa_1}\Big]+(\tau^c+\tau^{ss})y \tag{17}$$

I first calibrate the model using the baseline tax schedule and use it in the promise-keeping constraint when discussing optimal taxation and education policy.

The agents ability distribution is Pareto-Lognormal with parameters (m_1,m_2,λ).

The following parameterization is used to quantify the model. The only parameter that I need to calibrate is γ. It is determined by equalizing educational fee with median income. I set $T = 38$ at annual frequency.

There are two cohorts indexed respectively by σ_h and σ_l. The only place that aggregate shocks enter the economy is by changing the labor productivity $y_t = \sigma\theta z_t$. I choose a typical shock of 5%, consistent with empirical evidence in Oreopoulos et al. (2012).

Table 1: Model Parameters

Parameter	Value	Sources/Targets
β	0.96	Interest rate = 0.04
ν	0.25	Kapicka (2015)
α	0.5	Browning et al. (1999)
δ	0.0114	Hugget et al. (2011)
m_1	5.2483	Kapicka (2015)
m_2	0.4540	Kapicka (2015)
λ	2.0530	Kapicka (2015)
γ	0.15	Educational fee $=$ Average income
τ^c	0.052	Mendoza et al. (1994)
τ^{ss}	0.124	Kapicka (2015)
κ_0, κ_1	0.258, 0.768	Gouveia and Strauss (1994)
κ_2	0.0278	Kapicka (2015)
σ_h, σ_l	1.05,1	Oreopoulos et al. (2012)

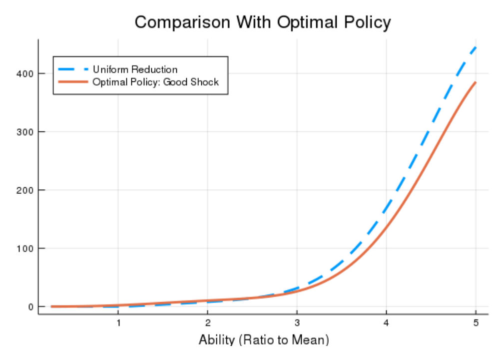
I consider two sets of policies. The first one is the baseline policy where the tax schedule is given by equation (17) and the educational fee is fixed. There are several interpretations of a fixed educational fee. One is that college tuition does not correlate with aggregate conditions. Another interpretation is related to student loan interest rate which also determines the cost of financing higher education. It is shown in Liu (2019) that student loan interest rate does not depend on aggregate labor market entry conditions. If educational fee is equalized across cohorts, agents in boom would choose higher level of education. To equalize educational resource across cohorts, I decrease the educational choice uniformly in boom for each type of agents.

The second set of policy is the optimal labor tax given by equation (9) and differential educational fees equation (10) indexed to labor market entry shocks. The promised utility is obtained from the baseline model. Importantly, the promised utility depends on labor market entry shock.

5.1 Optimal Education Resource Allocation

I first compare the educational choice distribution across cohorts. On the left of figure 1, it can be seen that under the optimal policy, the educational choice distributions closely resemble one another during boom and during bust. On the other hand, in the baseline policy with fixed educational fee and uniform educational choice reduction, it is clear from the right figure in figure 1 that the low ability agents experience a large reduction in their education levels while high ability agents gain more education access. Compared to the baseline, the higher educational fee in boom reduces inequality in educational choices. The calibration shows that the educational fee should be 9% larger when the cohort receives a 5% good aggregate shock, in order to equalize educational resources.

(a) Educational Choice: Optimal Policy



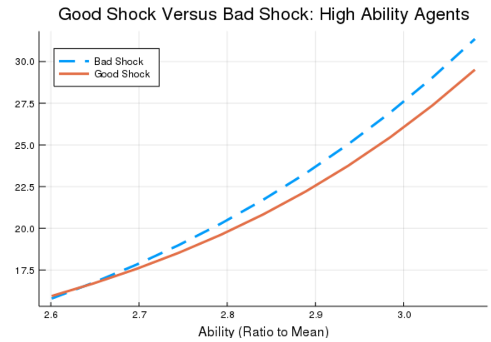
(b) Educational Choice: Baseline Policy

Figure 1: Educational Choice Across Cohorts

Although educational choice distributions look similar across cohorts, there are differences depending on agent types. Figure 2 zooms in on the left graph in figure 1 and compare educational choices across cohorts for different types of individuals.

What is interesting is that good shock benefits more the low ability agents as more educational resources are directed towards them compared to bust. On the other hand, high ability agents receive less educational resource in a boom than in a bust. This means when good shock hits, the social planner redistribute the limited educational resource so

(a) Educational Choice: Low Ability



(b) Educational Choice: High Ability

Figure 2: Educational Choice Across Cohorts: Zoom-in

that low ability agents get more education while high ability agents get less. To build intuition, let us refer back to equation (12) which now becomes:

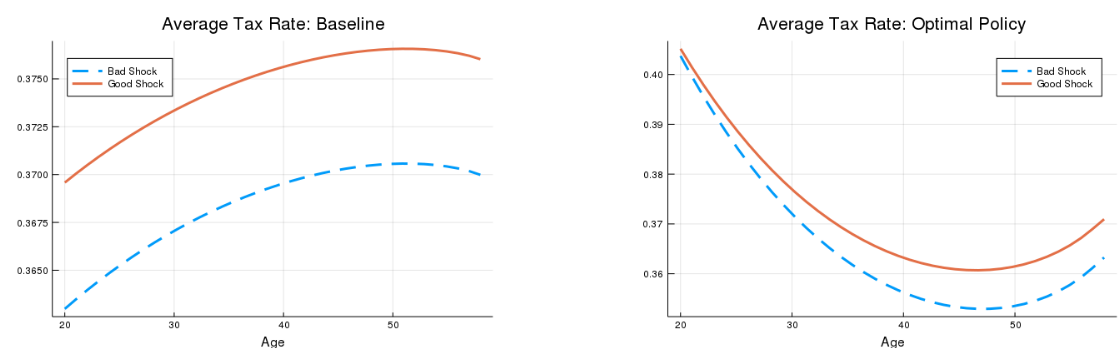
$$-\sum_{t=1}^{T} W_{z_t h_1} z_t X_\mu \gamma \log(\theta) \tag{18}$$

It can be shown that the other terms in equation (10) are negligible compared to equation (18), the marginal reduction in labor disutility. Its concavity of $W_{z_t h_1}$ in (z_t, h_1) is the main driving force of the result. For low ability workers, labor supply and educational choice are small, and the marginal disutility reduction $W_{z_t h_1}$ is quite sensitive to aggregate shocks. In other words, a good shock would lead to a large disutility reduction for low ability agents, hence higher marginal benefit. Even though low ability agents face a fee increase in boom, their marginal benefit from education increases more than the fee increase. In order to equalize benefit and cost, they choose higher educational levels which decreases marginal benefit. For high ability workers, the marginal disutility reduction is less sensitive to aggregate shocks due to higher level of labor supply and educational choice. If educational fee is equalized across cohorts, high ability workers would only increase their educational choice slightly when facing a good entry shock. The increase in educational fee more than offsets the small increase in educational choice. Instead, it causes high ability agents to choose lower levels of education.

5.2 Optimal Taxation

The optimal taxation takes into account the sum of own- and cross- Frisch-elascity ρ_t and educational tax deduction ϕ_t. With unobservable human capital, labor decisions are non-separable, so that cross-Frisch-elasticity can be positive or negative, depending on whether labor supply across periods are substitutes or complements.

I first show the average tax rate, weighted by the ability distribution $f(\theta)$, in figure 3. The left-hand-side plots the baseline average tax rate by age and the right-hand-side plots the optimal tax rate by age. The hump-shaped tax profile in the baseline arises from the hump-shaped labor supply over the life-cycle, and the gap between the boom- and bust-baseline tax rate remains almost constant through out the life-cycle. Since the baseline tax rate is progressive, this implies that the average-income gap across cohorts reduces as agents age, which is shown on the left-hand side in figure 5.



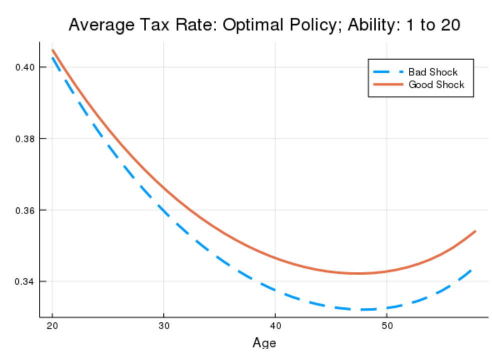
(a) Average Tax Rate: Baseline

(b) Average Tax Rate: Optimal Policy

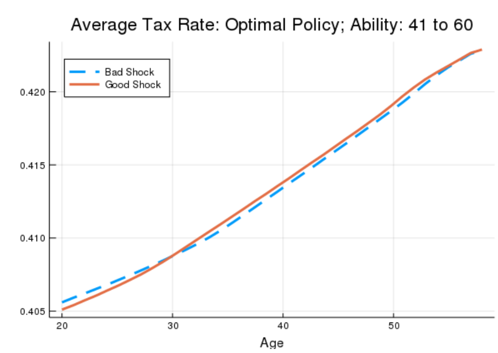
Figure 3: Average Tax Rates

On the other hand, the gap between optimal tax rates in boom and bust widens as agents age. This is due to negative cross-Frisch elasticity when agents just enter the labor market. In other words, labor supply in early periods is substitute for labor supply in later periods due to human capital accumulation. The social planner thus would like to set a high tax rate to discourage labor supply early on hence encouraging human capital accumulation. As agents age, the gain of human capital accumulation decreases, leading to declines in tax rates and widening of the tax gap.

How sensitive is cross-Frisch elasticity, or the pattern observed on the right-hand side in figure 3 to initial human capital, hence educational resource allocation? I divide the 100 ability grids into 5 equal size groups and show the optimal taxation comparison for boom and bust cohorts. Because education attainment is increasing in ability, I effectively group agents by initial human capital. The diminishing return to human capital accumulation is obvious: on the left-hand side of figure 4, the tax rates and the tax gap resembles those in figure 3, suggesting the latter is a result of the former. As initial human capital increases, on-the-job human capital accumulation becomes less important, weakening the connection between labor supply across periods. Hence, the optimal taxation the behaves more like the baseline taxation. This can be seen on the right-hand side in figure 4. Notably, the tax gap between boom and bust cohorts is minimal, suggesting that high initial human capital can ease the impact of labor market entry shock. This corresponds to the empirical evidence in Oreopoulos et al. (2012) who document that graduates from high-earning colleges and majors suffer less from graduating in recessions.



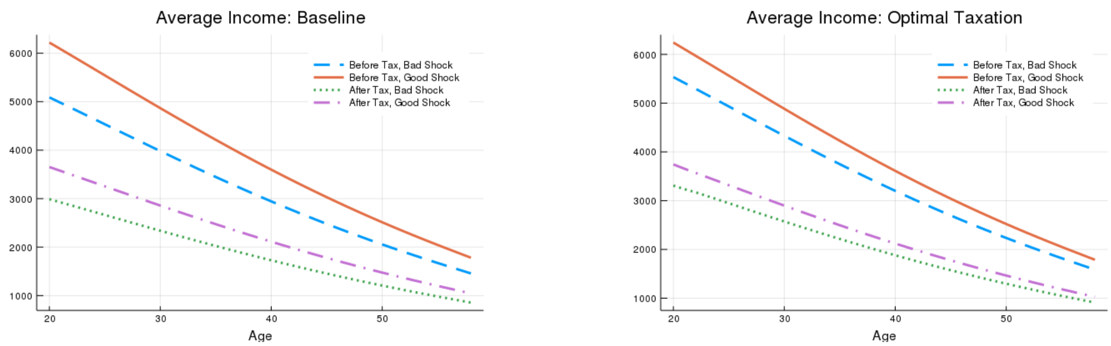
(a) Average Tax: Low Initial Human Capital



(b) Average Tax: High Initial Human Capital

Figure 4: Average Tax Rates: Low Versus High Initial Human Capital

Although the social planner does not equalize promised utility across cohorts, the income gap reduces under optimal taxation compared to the baseline. After tax income is 12% higher for boom cohort under the baseline taxation, while it is only 5.5% higher under the optimal taxation, shown in figure 5. The reasons are twofold: first, optimal taxation encourages human capital accumulation, and higher human capital reduces income gap due to labor market entry shock. Second, tax deduction due to education, i.e. ϕ_t is higher when initial human capital is low. Because low ability agents have less initial human capital in a bust than their counterparts in a boom, they receive more educational tax deduction, reducing the income gap across cohorts.



(a) Average Income: Baseline

(b) Average Income: Optimal Taxation

Figure 5: Average Income

How does educational tax reduction affect optimal taxation? I set ϕ_t to zero and compare the resulting tax schedule in figure 6. The gap between the optimal taxation and the one without deduction increases in later periods, indicating an increase in educational tax deduction. This is because higher initial human capital leads to more marginal labor disutility reduction in later periods. In other words, the benefit of higher initial human capital accumulates through out the life-cycle.

6 Implementation

In this section I discuss several implementations of the optimal policy in section 3. It can be shown that the optimal taxation can be implemented by a history-dependent tax system. However, in reality history-dependent tax system is very rare. [1] I discuss the welfare gain from several history-independent taxation and educational policies in this section.

[1]The optimal allocation can be implemented by history-independent tax system under very restrictive conditions. See Kapicka (2006).

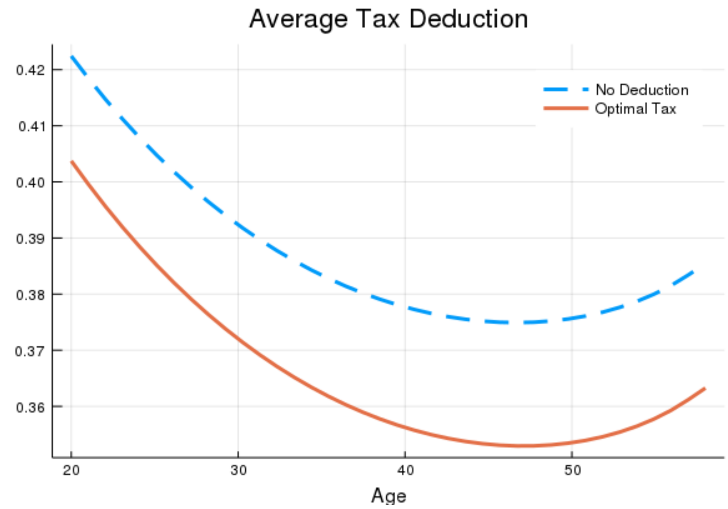


Figure 6: Average Educational Tax Deduction

6.1 A Static Income Tax System and an Income-Contingent Student Loan Design

Note first that even if the tax system cannot be made history-dependent, the cost-minimizing social planner would still consider a tax reform because she is aware of the productivity response from a change in taxation. This contrasts a naive social planner who would never change the tax system because under the current tax code, the allocations are already "optimal" to the social planner. I also assume that the tax system can be made age-dependent.

The way the social planner design a history-independent income tax system and educational fee policy is as follows: she first calculate the labor wedge of the cohort with bad entry shocks, at every period. Because the tax can be made age-dependent, if there is only one cohort, this tax system essentially implements the optimal taxation in section 3 if there is a one-to-one mapping between income at each period and ability. It is easy to show numerically that this is indeed the case.

However, the optimal taxation in section 3 takes into account the entry shock so that different cohorts face different tax rates. This is not feasible under a single income- and age-dependent tax system. However, with the help of educational policy, and specifically an income-contingent student loan repayment design, the social planner can implement the optimal policy under some conditions. Assumption 1 gives the conditions.

Assumption 1. *1. For agents with the same ability, there is a one-to-one mapping between optimal labor wedge and labor market entry shock.*

2. For agents with the same age, there is a one-to-one mapping between ability and income, taking entry shocks into consideration.

The first assumption can easily been shown under monotonicity of $y(\theta)$. The second one does not hold if the entry shock does not alter the educational choices. However, due to educational choice differences, the second assumption can numerically be shown to hold.

The social planner hence set the student loan repayment rate m_t to be such that

$$1 - \tau_t(\theta, \sigma_h) = (1 - \tau_t(\theta, \sigma_l))(1 - m_t)$$

That is, the effective tax rate for cohorts with good entry shock is the optimal labor wedge in lemma 1. Different from Stantcheva (2017), the age-dependent tax system plus income-contingent student loan repayment is able to implement the optimal policies because there is no human capital subsidy.

6.2 A Single Income-Contingent Student Loan Design

In reality, age-dependent tax system is also uncommon. Hence, I fix the tax schedule to be the one in the baseline and consider only the optimal educational fee policy implemented by income-contingent student loan repayment design. I also restrict the student loan design to be only income-contingent. That is, it cannot depend on age as in section 6.1.

Two policies are consider. In the first one, the income-contingent student loan repayment is fixed for all cohorts. Regardless of the entry shocks, agents with the same income pay the same rate of loan repayment. The loan repayment lasts for 20 periods. The way the loan repayment rate is calculated is as follows: the social planner first observes both in the baseline case and in the optimal policy the income distribution for all agents who enter the economy with 20 periods or less. She then calculate for each income level, the average tax rate difference between the baseline case and the optimal policy. She then sets the student loan repayment rate such that each income level, the average tax rate is equalized. The agents then choose their educational choice and labor supply under the new policies.

The result shows that the loan repayment rate is 3.4% and the welfare gain is 0.2% which is around 25% of the total welfare gain under the optimal policy. These findings are consistent with Findeisen and Sachs (2016) who find that income-contingent repayment system can achieve welfare gains from 0.2% to 0.6% of lifetime consumption. The intuition of the result is straight-forward: the income-contingent student loan effectively change the implied tax rate so that the allocations are closer to the optimal.

I consider a second case in which the loan repayment rate can differ across cohorts. This can be implemented by an income-contingent student loan repayment plan plus a student loan interest rate indexed to labor market entry shocks. The calculation of the loan repayment rate is similar.

For the cohort with 5% good entry shock, the loan repayment rate is 3.9% while the repayment rate for the unlucky cohort is 3.1%. The welfare gain increases to 0.33% or 40% of the total welfare gain of the optimal policies.

This result emphasizes the need to take into consideration labor market entry shocks in student loan design. Specifically, the current student loan interest rates and repayment rates are determined at the time of loan receipt and do not correlate with aggregate labor market conditions. Thus, cohorts that enter the labor market in a recession are likely to face more difficulties repaying their loans, have higher rate of default, and result in worse labor market outcomes.

A simple regression shows the lack of correlation between student loan interest rates and aggregate labor market conditions. Table 2 regress the interest rates of the federal direct unsubsidized loans on unemployment rates or future unemployment rates. [2] The result shows that there is no correlation between the interest rates and contemporaneous unemployment rates, nor is there correlation with future unemployment rates. This implies that the cost of financing higher education via student loans is disconnected with the potential gain from educational investment.

Table 2: Federal Student Loan Interest Rates and Unemployment

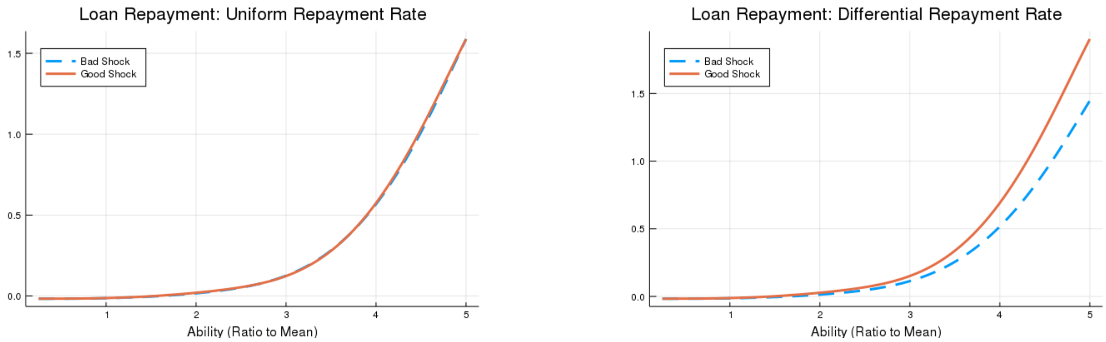
	(1)	(2)	(3)	(4)	(5)
Unemployment rate	0.0194				
	(0.140)				
F.unemployment rate		0.118			
		(0.130)			
F2.unemployment rate			0.202		
			(0.143)		
F3.unemployment rate				0.182	
				(0.150)	
F4.unemployment rate					0.041
					(0.156)
Observations	27	26	25	24	23

The interest rates are on the federal direct unsubsidized loan. F is the lead operator. Hence F.unemployment rate means unemployment rate one year from current observation.

I compare the loan repayment distribution in both cases. For each ability level, I calculate the total loan repayment and plot the distribution in figure 7. According to the plot on the left-hand side of figure 7, the total repayment is quite similar across cohorts cohorts, suggesting that the income-contingent student loan is able to partially reduce the income inequality across cohorts. On the right-hand side of the figure, the total loan repayment is much higher for high ability workers with good entry shock.

Because loan outstanding is forgiven after 20 periods, the system suggests that low ability agents pay less than the total educational fee while high ability agents pay more

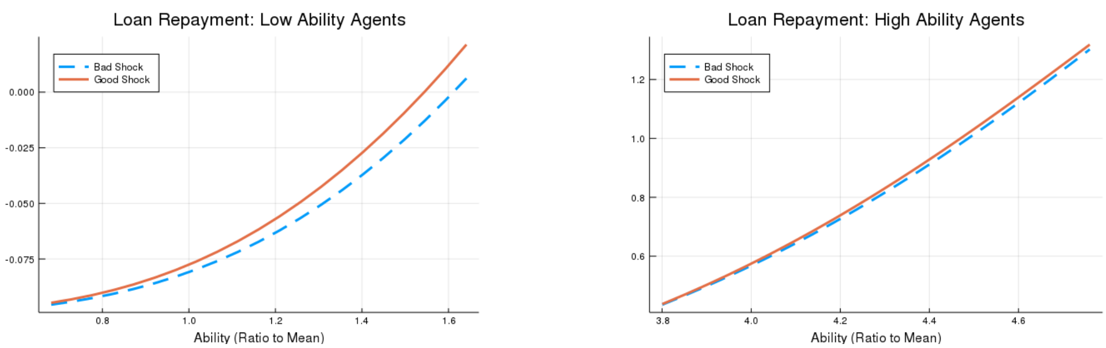
[2] The federal direct loan is the large student loan program. Over 90% of current outstanding student loan is federal loan.



(a) Uniform Repayment Rate (b) Differential Repayment Rate

Figure 7: Loan Repayment Distributions

than the total educational fee. Findeisen and Sachs (2016) also emphasize this point. I illustrate the result in figure 8. The figure shows the total loan repayment as a fraction of the educational fee when there is a uniform repayment rate, separately for low ability and high ability agents. In other words, they are zoomed-in version of the left plot in figure 7.



(a) Low Ability Agents' Loan Repayment (b) High Ability Agents' Loan Repayment

Figure 8: High and Low Ability Agents' Loan Repayment

For low ability workers, the loan forgiveness percentage could be as much as 7.5% while the high ability workers can pay up to 1.2 times the educational fee in total loan repayment. Both low and high ability workers with good entry shock pay more than their counterpart with bad entry shock. However, the percentage difference is larger among low ability workers, suggesting that labor supply is more sensitive to entry shocks for low ability workers than for high ability workers.

7 Conclusion

I study optimal taxation and educational policy in an environment in which with limited educational resource and observable labor market entry shocks that persist through life-cycle. Human capital accumulation is unobservable which makes labor decision non-separable. In the optimal policies, the educational resource allocation varies with entry shocks, with low ability agents receiving more while high ability agents receiving less with a good entry shock. The optimal labor wedge decreases in early periods when initial human capital is low to incentivize human capital accumulation. The pattern reverses when initial human capital is high. I calibrate the model to the U.S. data and show that the overall welfare gain is 0.8%. I show how to implement the optimal taxation with an age-dependent labor tax system and an income-contingent student loan repayment design with interest rates indexed to labor market entry shocks. A single income-contingent student loan with interest rate adjustment can achieve 40% of the welfare gain. The intuition is that income-contingent student loan repayment effectively changes implied tax rate so that allocations are closer to the optimal.

The result emphasizes the need to take into consideration labor market entry shocks in student loan design. Specifically, the current student loan interest rates and repayment rates are determined at the time of loan receipt and do not correlate with aggregate labor market conditions. Thus, cohorts that enter the labor market in a recession are likely to face more difficulties repaying their loans, have higher rate of default, and result in worse labor market outcomes. I also show that in the single income-contingent student loan design, it is likely that low ability agents pay less than the total loan amount while high ability agents pay more than the total loan amount.

References

Altonji, Joseph G., Lisa B. Kahn, and Jamin D. Speer, "Cashier or Consultant? Entry

Labor Market Conditions, Field of Study, and Career Success," *Journal of Labor Economics*, 2016, *34* (51), 361–401.

Browning, Martin, Lars Peter Hansen, and James J. Heckman, "Micro Data and General Equilibrium Models," *Handbook of Macroeconomics*, 1999, *1*, 543–633.

Findeisen, Sebastian and Dominik Sachs, "Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans," *Journal of Public Economics*, 2016, *138*, 1–21.

Gouveia, Miguel and Robert P. Strauss, "Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis," *National Tax Journal*, 1994, *47* (2), 317–339.

Hugget, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *American Economic Review*, 2011, *101* (7), 2923–2954.

Kahn, Lisa B., "The Long-Term Labor Market Consequences of Graduating From College in a Bad Economy," *Labour Economics*, 2010, *17* (2), 303–316.

Kapicka, Marek, "Optimal Income Taxation with Human Capital Accumulation and Limited Record Keeping," *Review of Economic Dynamics*, 2006, *9* (4), 612–639.

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

Liu, Kai, Kjell G. Salvanes, and Erik Sorensen, "Good Skills in Bad Times: Cyclical Skill Mismatch and the Long-Term Effects of Graduating in a Recession," *European Economic Review*, 2016, *84*, 3–17.

Liu, Yizhou Andrew, "Optimal Student Loan Design and Human Capital Policy with Labor Market Entry Shock," *Working Paper*, 2019.

Mendoza, Enrique G., Assaf Razin, and Linda L. Tesar, "Effective Tax Rates in Macroeconomics: Cross-country Estimates of Tax Rates on Factor Incomes and Consumption," *Journal of Monetary Economics*, 1994, *34* (3), 297–323.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz, "The Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, 2012, *4* (1), 1–29.

Stantcheva, Stefanie, "Optimal Taxation and Human Capital Policies over the Life Cycle," *Journal of Political Economy*, 2017, *125* (6), 1931–1988.