Human Capital Investment and Optimal Income Taxes over the Life Cycle[†]

Been-Lon Chen[‡]
Academia Sinica

Fei-Chi Liang[*]
National Taiwan University

February 2017

Abstract

We study a life-cycle model with heterogeneous agents of discrete skill types. In the model, unobservable skills evolve over time through endogenous human capital investment, rather than via stochastic shocks. Our main findings are as follows. First, even though our model has no uncertainty and thus no insurance motive, the capital wedge is positive. Next, the labor wedge is neither always positive nor constant over time, but is negative in first period and ambiguous before the terminal period of the life cycle. Finally, these wedges can be implemented as linear taxes on capital and labor, along with lump-sum taxes, in the competitive market and there is a welfare gain from the second-best optimal mechanism, with the gain increasing in the gap of agents' skills.

1. Introduction

This paper studies how human capital investment affects the design of optimal income tax policies. We consider a life-cycle model with discrete skill-type agents whose abilities to work are augmented by unobservable human capital investment. In the existing dynamic Mirrlees literature, people differ simply through their skills which follow a stochastic process over time. The government wants to redistribute from high-skill to low-skill agents but can only observe earnings (and not abilities), leading to a non-degenerate equity-efficiency trade off. Thus, the optimal tax system is designed based on insurance and redistribution motives. Along the literature, if the skill distribution is fixed and thus there is no uncertainty on skills, a zero capital wedge is called for, because of the lack of motives to insure against lifetime risks. (e.g., Werning, 2007; da Costa and Masestri, 2007; Anderberg, 2009; Farhi and Werning 2013; Kopczuk, 2013; Stantcheva, 2016). Moreover, the labor wedge for low-skill types is positive, so that high-skill agents do not mimic low-skill types and work less (e.g., Stiglitz, 1982; Golosov et al. 2006; Piketty and Saez, 2013). In particular, the labor wedge is constant over time (perfect labor tax smoothing), unless there are persistent idiosyncratic shocks on skills that vary with aggregate shocks (e.g., Golosov et al. 2006; Werning, 2007; Farhi and Werning, 2013). The departure of our paper from the dynamic Mirrlees literature is that we study optimal income tax policies when skills evolve according to unobservable human capital investment rather than stochastic uncertainties.

We use a discrete skill-type model. To simplify the analysis, we follow Stiglitz (1982) and posit two types of agents, differing in abilities to acquire skills but having the same utility function. In addition to working and savings, all agents choose human capital investment. Agents' heterogeneities in skills mainly come from endogenous human capital investment. To streamline the study, we assume that when born, both types of agents have identical human capital levels and thus identical skills, but the high-skill type has advantages in accumulating skills.[1] Expenses for human capital investment is non-verifiable: private expenditures for consumption may be pretended as private expenses for education purposes and are not distinguishable from the viewpoint of the government.[2] Under asymmetric information, the government (the social planner) solves the second-best program: it chooses the (constrained) optimal allocations to maximize the utilitarian social welfare subject to resource constraints and incentive compatibility constraints.

---

[1] Such initial advantages to accumulate human capital capture innate abilities and cognitive and noncognitive skills in early childhood development, as emphasized by Todd and Wolpin (2003, 2007) and Cunha and Heckman (2008)

[2] For non-verifiable investment in human capital, see, among others, Bovenberg and Jacobs (2005), Kapička (2006, 2015), and Grochulski and Piskorski (2010). For example, Bovenberg and Jacobs (2005) argued that, in practice, books, computers and travelling costs are difficult to verify, because individuals may misrepresent expenditures for private consumption purposes as expenses for education investment.

We obtain two novel findings concerning the constrained efficient allocations. First, even though our model does not have any uncertainties and thus no role for insurance purposes, the capital wedge on low skills is positive. Moreover, the labor wedge on low skills is neither always positive nor constant over time, but it is positive in the terminal period and negative in first period and ambiguous in all other periods of the life cycle.

These wedges arise, because human capital investment is endogenous and non-verifiable by the government. These wedges aim to foster human capital investment. First, with unobservable human capital investment, if high-skill agents shirk, the benefit is not only from working less for leisure, but also from reducing expenses on education for more consumption. Thus, even without uncertainty on skills, the intertemporal marginal rate of substitution in consumption is distorted by the informational friction concerning human capital investment. A positive capital wedge on low skills is optimal, because the policy discourages high-skill agents from misreporting low skills and from reducing unobservable expenses in human capital investment. Next, a negative labor wedge on low skills in the first period and possibly in other early periods of the life cycle is optimal, because the policy attracts low-skill agents to work more early in their life-cycle. The policy deters high-skill agents from misreporting as low skills; if they misreport as low skills, they have to work more.

While it is tempting to interpret these capital and labor wedges as actual taxes on capital and labor, the relationship between wedges and taxes is not straightforward, because there is a *double deviation* problem.[3] The tax implementation is to find tax systems so that the resulting competitive equilibrium yields these optimal allocations. This paper proposes a history-dependent tax system, wherein capital and labor income are taxed linearly, along with lump-sum taxes, if an agent's history of capital and effective labor satisfies some conditions; otherwise, an agent would face extremely high taxes. We show that, under this tax system, the linear tax rates are consistent with the optimal capital and labor wedges.

Finally, we carry out numerical analysis. We find that the consideration of endogenous human capital increases capital wedges and decreases labor wedges. Moreover, there is a welfare gain from our second-best optimal mechanism relative to the laissez-fare economy with linear taxes, with the welfare gain increasing in the gap of abilities between agents.

---

[3] Intuitively, each wedge controls only one aspect of worker's behavior (labor in a period, or savings) taking all other choices fixed *at the optimal level*. For example, assuming that an agent supplies the socially optimal amount of labor, a capital tax defined by an intertemporal wedge would ensure that the agent also makes a socially optimal amount of savings. However, agents choose labor and savings jointly; if an agent considers to change her labor, then, in general, she also considers to change her savings. Thus, there are *double deviations*. Kocherlakota (2005), Albanesi and Sleet (2006) and Golosov and Tsyvinski (2006) showed that such double deviations would give an agent a higher utility than the utility from the socially optimal allocations, and therefore the optimal tax system must be enriched with additional elements in order to implement the optimal allocations.

Related literature

Our paper is related to human capital accumulation and the optimal taxation. The process of human capital formation has been a long-lasting literature, starting with Becker (1964), Ben-Porth (1967) and Heckman (1976). The structural branch of the literature emphasizes that human capital acquisition occurs throughout a life cycle, underscoring the need for a life cycle model (Cunha and Heckman, 2007). Ex ante heterogeneity in the returns to human capital matters. A large body of empirical work documents the importance of human capital as a determinant of earnings (Goldein and Katz, 2008), and cognitive and noncognitive skills as being equally important (Todd, P. and Wolpin, 2003; Cunha and Heckman, 2008). The model developed in our paper attempts to embrace some of this literature's main findings in a stylized way.

There is a growing literature named new dynamic public finance which analyzes the optimal taxation pioneered by Mirrlees (1971) in dynamic settings. As opposed to the Ramsey approach wherein agents are homogeneous and information is complete, agents are heterogeneous in earning skills that are private information in the Mirrlees approach. In the Mirrlees framework, the benevolent government chooses the allocation that trades off between efficiency and equity. The new dynamic public finance literature typically considers exogenously evolving abilities, thus abstracting from endogenous skill acquisition.[4] Our paper contributes to this literature by taking into account individuals' skills which evolve over time based on endogenous human capital investment.

A series of papers in the dynamic Mirrlees approach have jointly considered optimal taxation and endogenous human capital.[5] Investment in human capital may take the form of labor effort and expenditures. Thus, the existing model can be divided into two strands. Though different, our paper uses expenses as investment in human capital and is complementary to the strand that uses labor as input.

In our paper, expenses for human capital investment are non-verifiable. In a static model, Bovenberg and Jacobs (2005) considered both verifiable and non-verifiable expenses for human capital investment. They found positive optimal income taxes for re-distributional purposes and positive optimal subsidies on verifiable education expenses for offsetting some tax-induced distortions on learning.[6] In a dynamic

---

[4] See Golosov et al. (2003), Kocherlakota (2005), Albanesi and Sleet (2006), Werning (2007), Farhi et al. (2012), and Farhi and Werning (2013), among others. Golosov et al. (2006), Kocherlakota (2010), Piketty and Saez (2013) and Kopczuk (2013) provided excellent survey.

[5] See Kapička (2006, 2015), da Costa and Masestri (2007), Boháček and Kapička (2008), Anderberg (2009) and Stantcheva (2015). Among these papers, education time is unobservable in Kapička (2006), Anderberg (2009) and Kapička (2015) and is observable in other studies.

[6] For other static models, see also Maldonado (2007), Jacobs and Bovenberg (2011), DaCosta and Maestri (2007), Gelber and Weinzierl (2012) and Findeisen and Sachs (2016).

3

model with only verifiable education expenses, Stantcheva (2016) highlighted the importance of the complementarity between ability and education, which can be used to measure the distortion to human capital. Moreover, in these two papers, agents' earning skills are also affected by stochastic shocks, which lead to a positive capital tax. Our model is different from these two existing papers in that we study only non-verifiable education expenses and moreover agents' earning skills are not affected by stochastic shocks. Yet, even without uncertainties on skills, it is optimal to tax capital income on low types in our model. In particular, our capital taxation on low types serves as a mechanism to increase educational investment, as opposed to educational subsidies proposed by these two papers.

Kapička (2006, 2015) and Grochulski and Piskorski (2010) also study dynamic models with the setting of unobservable human capital investment, so there are no feasible schooling policies. First, our model assumes ex ante different skill types, as opposed to ex ante identical agents with ex post different skill types in Grochulski and Piskorski (2010). Next, in Kapička (2006, 2015), investment in human capital is labor time, wherein different skill types do not affect human capital formation. By contrast, in our model, investment in human capital is expenses, wherein ex ante different skill types affect human capital formation. In Grochulski and Piskorski (2010), ex ante identical agents invest in human capital only in the initial period, and then, agents' human capital may completely depreciate due to stochastic depreciation shocks, so some agents enter a low human capital state, which is an absorbing state. As a result, their labor wedge is always positive for low skills and, except in the terminal period of an agent's life, is always negative for high skills. Note that, in the case without stochastic shocks on skills in Grochulski and Piskorski (2010), the capital wedge is zero in every period. In contrast, our model has ex ante heterogeneous agents, who invest in human capital in all except the terminal period of their life. Hence, even without stochastic shocks on skills, the capital wedge is positive in all except the initial period when all agents are born with the same physical capital level. Besides, because high and low skill agents invest in human capital, the labor wedge for low skills is negative in the first period and may be negative or positive in all other periods except the final period.

On the technical side, several papers studied models with agents of a continuous distribution of skills (e.g., Farhi et al., 2012; Farhi and Werning, 2013; Kapička and Neira, 2015), dubbed as the first-order approach, since their incentive compatible constraints are typically written in terms of envelop conditions. As these envelop conditions are only necessary but not sufficient, the solution to the program might not be a solution to the full program (Ebert, 1992). Thus, the approach needs to validate that the constrained efficient allocations solved by these conditions indeed give the utility intended by the planner (e.g., Farhi and Werning, 2013; Stantcheva, 2016). Our model posits skills of a discrete type, and the solutions are necessary and sufficient. Moreover, even if the first-order approach is used, in the Appendix

we have shown that, except for the top and bottom ability in the distribution, our results continue to hold.

Finally, to tackle double deviation problems, Albanesi and Sleet (2006) implemented the constrained efficient allocations in terms of non-linear taxes in a competitive equilibrium. They showed that these taxes are non-separable in wealth and labor and depend in each period on agents' wealth and labor income in that period and not on other aspects and past history. By restricting to linear capital taxes and arbitrarily nonlinear labor income taxes, Kocherlakota (2005) implemented the constrained efficient allocations by separating capital from labor taxes and both taxes are history-dependent. Following the tax structure in Kocherlakota (2005), Grochulski and Piskorski (2010) found that deferred capital taxes are the necessary condition for linear capital taxes, with negative expected capital taxes early in the life-cycle and positive expected capital taxes later in the life-cycle so that the ex ante expected present value of lifetime capital taxes is zero. Parallel to these studies, our paper proposes a non-separable and history-dependent tax system to implement the constrained efficient allocations. We show that the optimal linear tax on capital and labor income in this tax system are exactly the capital and labor wedges.

We organize this paper as follows. In Section 2, we present the model. The social planner's problem is studied in Section 3, and the signs of capital and labor wedges are analyzed in Section 4. Section 5 provides a tax system to implement the constrained efficient allocation obtained in the planner's problem as a competitive equilibrium. In Section 6, we offer numerical analysis. Finally, concluding remarks are offered in Section 7.

2.   Basic Model

2.1   The Environment

The economy consists of a continuum of agents who live for T years, during which they work, consume, and invest in physical capital and human capital. There are two types of agents, the high-skilled and the low-skilled, denoted by H and L, respectively, with the former accounting for the fraction π^H and thus, the latter for the remaining fraction $\pi^L \equiv 1 - \pi^H$. We will also refer to the high-skilled as high skills or high types, and the low-skilled as low skills or low types.

An agent's preference is represented by the following lifetime utility function:

$$\sum_{t=1}^{T} \beta^{t-1} \big[u(c_t) - v(l_t) \big],$$

where $0 < \beta < 1$ is the discount factor, c_t is consumption in period t and l_t is work effort in period t. An agent at most provides $\bar{l} > 0$ work effort in a period. We assume that $u(c)$ is continuously differentiable, strictly increasing and concave, and satisfies the Inada condition, and $v(l)$ is continuously differentiable,

strictly increasing and convex, and satisfies $v(0)=0$, $\lim_{l\to 0} v'(l)=0$ and $\lim_{l\to \bar{l}} v'(l)=\infty$.

In period t, an agent's disposable income may be consumed, spent to accumulate human capital h_{t+1}, or saved to form physical capital k_{t+1}. Human capital characterizes working skill levels: an agent with human capital h_t and work effort l_t supplies $z_t=l_t h_t$ effective labor. The human capital technology is $h_{t+1}^i = \psi_t^i(h_t^i, e_t^i)$, $i=H, L$, where e_t^i is educational expenses. When born, both types of agents are endowed with identical human capital levels $(h_1^H = h_1^L)$, but type H accumulates human capital faster than type L; namely, for given levels of h and e, $\psi_t^H(h,e) \geq \psi_t^L(h,e)$.[7]

In period t, the representative firm combines aggregate physical capital K_t and aggregate effective labor Z_t to produce final goods using the technology $Y_t = F(K_t, Z_t)$. The technology is neoclassical which satisfies constant returns and is strictly increasing and concave in both arguments. The physical capital depreciates at the rate of δ_k.

In our environment, individual physical capital k_t and effective labor z_t are publicly observable, while individual consumption c_t, human capital h_t, educational expenses e_t, work effort l_t and agents' types are not publicly observable. However, the sum of an agent's consumption and human capital spending is observable, as it is inferable from capital and labor income which are observable. Moreover, since agents have no incentives to invest in human capital in the terminal period, actual consumption c_T is observable in the terminal period.

2.2 Incentive Compatibility

Without incentive compatibility constraints, in order to maximize the social welfare, the social planner will ask high-skill agents to provide more effective labor but allocates consumption equally to each type. However, work effort and educational expenses are private information. Allocation of equal consumption for both types encourages the high-skilled to mimic the low-skilled and reduces their work effort and educational expenses. It is worth noting that only high-skill agents have incentives to deviate, because under equal consumption allocation for both types, low-skill agents do not gain if they misreport as high skills. By the revelation principle, we focus on the direct revelation mechanism.

Agents of type $i \in \{H, L\}$ report their types. Since agents can either truthfully report or lie about their types, there are four patterns of strategies. Following da Costa and Maestri (2007) and Farhi and Werning (2013), we denote $\sigma=(r|i)$ as a strategy of reporting type r given true type i. As only the high-

---

[7] In a paper with two skill types and with time spent in human capital formation, da Costa and Maestri (2007) also assume that, with other things being equal, high types accumulate human capital faster than low types.

skilled have incentives to misreport their types, only the following three strategies are plausible: σ∈{$H|H$, $L|L$, $L|H$}. If x is observable, notation x^i denotes what the social planner allocates to type i. If it is unobservable, then $x^{r|i}$ denotes the optimal choice by type i who reports as type r.

An allocation $A=(c, e, l, h, z, k, Z, K)$ specifies consumption $c=\{c_t^{H|H}, c_t^{L|L}, c_t^{L|H}\}_{t=1}^T$, educational expenses $e=\{e_t^{H|H}, e_t^{L|L}, e_t^{L|H}\}_{t=1}^T$, work effect $l=\{l_t^{H|H}, l_t^{L|L}, l_t^{L|H}\}_{t=1}^T$, human capital $h=\{h_t^{H|H}, h_t^{L|L}, h_t^{L|H}\}_{t=1}^T$, effective labor $z=\{z_t^H, z_t^L\}_{t=1}^T$, physical capital $k=\{k_t^H, k_t^L\}_{t=1}^T$, aggregate effective labor $Z=\{Z_t\}_{t=1}^T$ and aggregate physical capital $K=\{K_t\}_{t=1}^T$, given initial $h_1^{H|H}=h_1^{L|L}=h_1^{L|H}=h_1$ and $k_1^H=k_1^L=k_1=K_1$. Denote by G the government expenditure. Given $K_1=k_1$, h_1 and $G=\{G_t\}_{t=1}^T$, an allocation A is **resource feasible** if

$$\sum_{i=H,L} \pi^i \left[c_t^{i|i} + e_t^{i|i} + k_{t+1}^i \right] \le F\left(K_t, Z_t\right) + \left(1-\delta_k\right)K_t - G_t,$$

$$c_t^{r|i} + e_t^{r|i} = c_t^{r|r} + e_t^{r|r} \quad,$$

$$h_{t+1}^{r|i} = \psi_t^i\left(h_t^{r|i}, e_t^{r|i}\right) \quad \text{and} \quad l_t^{r|i} = \frac{z_t^i}{h_t^{r|i}},$$

for any $t\in\{1,2.. T\}$ and any $(r|i)\in\{H|H, L|L, L|H\}$, where $\sum_{i=H,L} \pi^i k_t^i = K_t$, $\sum_{i=H,L} \pi^i z_t^i = Z_t$,

We remark that the sum of an agent's consumption and educational expenses is observable by the public. Thus, agents with strategy $\sigma = r|i$ need to restrict the sum of their consumption and educational expenses $c_t^{r|i} + e_t^{r|i}$ to their reporting level $c_t^{r|r} + e_t^{r|r}$.

We focus on the allocation in which high-skill agents are truth-telling. An allocation A is **incentive compatible** (IC) if

$$\sum_{t=1}^T \beta^{t-1}\left[u\left(c_t^{H|H}\right) - v\left(\tfrac{z_t^H}{h_t^{H|H}}\right)\right] \ge \sum_{t=1}^T \beta^{t-1}\left[u\left(c_t^{L|H}\right) - v\left(\tfrac{z_t^L}{h_t^{L|H}}\right)\right],$$

where $c_t^{L|H}$ and $h_t^{L|H}$ are chosen to maximize the utility of high-skill deviators.

In the incentive compatible allocation, a truth-telling strategy ($H|H$) gives the utility larger than or equal to a misreporting strategy ($L|H$). We remark that there is no incentive compatibility constraint on low-skill agents, since even with informational frictions, the low-skilled do not benefit from misreporting as high skills. Hence, it is not necessary to prevent low-skill agents to deviate from their true types.

3. The Constrained Efficient Allocation

This section envisages the constrained efficient problem of the social planner. We focus on

7

maximizing a utilitarian social welfare function.[8] By deriving the second-best solution to this problem and the optimal allocation to the agent's problem, we can sign capital wedges and labor wedges.

The following definition describes the second-best optimal solution of the economy.

Definition 1. An allocation A^* is **constrained efficient** if it maximizes the welfare of the utilitarian social planner in the class of all feasible incentive compatible allocations.

The incentive compatible optimal allocation will be referred to as the constrained efficient allocation. Because of incentive compatibility constraints, the constrained efficient allocation does not satisfy the standard Euler equation in this model. This makes room for the benevolent government to impose the optimal tax to replicate the constrained efficient allocation.

3.1   The Social Planning Problem

The planner can observe an agent's physical capital k_t^i and effective labor z_t^i and, from an agent's budget constraint, can infer the sum of consumption and educational expenses $c_t^i + e_t^i$. The informational problem is that the planner cannot observe an agent's type and thus, observes neither an agent's consumption c_t^i and educational expenses e_t^i, nor an agent's human capital h_t^i. While the planner knows an agent's effective labor, she cannot know work effort $l_t^i = \frac{z_t^i}{h_t^i}$, since h_t^i is unobservable. Thus, if a high-skilled agent misreports as a low skill (i.e., $L|H$, to avoid being caught she must restrict the sum of the expenditure on consumption and educational expenses equal to the sum of the low-skilled: $c_t^{L|H} + e_t^{L|H} = c_t^{L|L} + e_t^{L|L}$.

The social planner chooses observable allocations $\{c_t^{i|i} + e_t^{i|i}, k_t^i, z_t^i\}_{t=1}^T$, $i \in \{H, L\}$, to maximize the following utilitarian social welfare problem:

$$\max \sum_{i=H,L} \pi^i \sum_{t=1}^T \beta^{t-1} \left[u\left(c_t^{i|i}\right) - v\left(\frac{z_t^i}{h_t^{i|i}}\right) \right], \tag{1a}$$

subject to the resource constraints for $t=1, 2, …, T$:

$$\sum_{i=H,L} \pi^i \left[c_t^{i|i} + e_t^{i|i} + k_{t+1}^i \right] \le F\left(\sum_{i=H,L} \pi^i k_t^i, \sum_{i=H,L} \pi^i z_t^i \right) + \left(1-\delta_k\right) \sum_{i=H,L} \pi^i k_t^i - G_t, \quad k_{T+1}^i = 0, \tag{1b}$$

and the incentive compatibility (henceforth, IC) constraint:

---

[8]   See Diamond (1998) and Tuomala (1990) concerning to how the choice of the welfare function affects optimal taxes in static framework.

$$\sum_{t=1}^{T} \beta^{t-1} \left[u\left(c_t^{H|H}\right) - v\left(\frac{z_t^H}{h_t^{H|H}}\right) \right] \geq \sum_{t=1}^{T} \beta^{t-1} \left[u\left(c_t^{L|L} + e_t^{L|L} - e_t^{L|H}\right) - v\left(\frac{z_t^L}{h_t^{L|H}}\right) \right], \tag{1c}$$

where condition $c_t^{L|H} = c_t^{L|L} + e_t^{L|L} - e_t^{L|H}$ is used in the right-hand side of (1c).

Let λ_t be the shadow price of the resource constraint at time t and μ be the shadow price of the IC constraint. Denoting the optimal allocation by one with an asterisk, then the first-order conditions for total spending $c_t^{i|i} + e_t^{i|i}$, effective labor z_t^i and capital k_t^i, $i \in \{H, L\}$, are as follows.

$$\beta^{t-1} u'\left(c_t^{*H|H}\right) \left[1 + \frac{\mu}{\pi^H} \right] = \lambda_t; \tag{2a}$$

$$\beta^{t-1} u'\left(c_t^{*L|L}\right) \left[1 - \frac{\mu}{\pi^L} \frac{u'\left(c_t^{*L|H}\right)}{u'\left(c_t^{*L|L}\right)} \right] = \lambda_t; \tag{2b}$$

$$\beta^{t-1} v'\left(\frac{z_t^{*H}}{h_t^{*H|H}}\right) \frac{1}{h_t^{*H|H}} \left[1 + \frac{\mu}{\pi^H} \right] = \lambda_t F_z\left(K_t^*, Z_t^*\right); \tag{2c}$$

$$\beta^{t-1} v'\left(\frac{z_t^{*L}}{h_t^{*L|L}}\right) \frac{1}{h_t^{*L|L}} \left[1 - \frac{\mu}{\pi^L} \frac{v'\left(\frac{z_t^{*L}}{h_t^{*L|H}}\right) \frac{1}{h_t^{*L|H}}}{v'\left(\frac{z_t^{*L}}{h_t^{*L|L}}\right) \frac{1}{h_t^{*L|L}}} \right] = \lambda_t F_z\left(K_t^*, Z_t^*\right); \tag{2d}$$

$$\lambda_t = \lambda_{t+1} \left[F_k\left(K_{t+1}^*, Z_{t+1}^*\right) + \left(1 - \delta_k\right) \right]. \tag{2e}$$

Condition (2e) is standard for capital accumulation. Conditions (2a)-(2d) reduce to the standard forms for total spending and effective labor, if the IC constraint is not binging and thus $\mu = 0$, in which case the discounted marginal utility of total spending and effective labor for each type is equal to their marginal cost. With a binding IC constraint and thus $\mu > 0$, the shadow price of the IC constraint augments the discounted gross marginal utility of consumption and labor for high types but detracts from that for low types.

Conditions (2a), (2c) and (2e) give the Euler equation and the tradeoff between total spending and effective labor for one who reports as a high type.

$$\frac{u'\left(c_t^{*H|H}\right)}{\beta u'\left(c_{t+1}^{*H|H}\right)} = \left[F_k\left(K_{t+1}^*, Z_{t+1}^*\right) + \left(1 - \delta_k\right) \right], \tag{3a}$$

$$v'\left(\frac{z_t^{*H}}{h_t^{*H|H}}\right) \frac{1}{h_t^{*H|H}} \frac{1}{u'\left(c_t^{*H|H}\right)} = F_z\left(K_t^*, Z_t^*\right). \tag{3b}$$

It is clear that in (3a), for high-type reporters, the social planner equalizes the marginal rate of substitution (henceforth, MRS) between consumption today and tomorrow to the marginal rate of transformation (henceforth, MRT) between consumption and investment today. As investment today accumulates capital tomorrow, the MRT between consumption and investment today is the marginal product of capital tomorrow. In (3b), for high-type reporters, the social planner equalize the MRS

9

between consumption and leisure today to the **MRT** between consumption and labor today. The MRT between consumption and labor today is the marginal product of labor today.

Similarly, for someone who reports as a low type, (2b), (2d) and (2e) imply the following relationships.

$$u'\left(c_t^{*L|L}\right)\left[1-\frac{\mu}{\pi^L}\frac{u'\left(c_t^{*L|H}\right)}{u'\left(c_t^{*L|L}\right)}\right]\frac{1}{\beta u'\left(c_{t+1}^{*L|L}\right)}\left[1-\frac{\mu}{\pi^L}\frac{u'\left(c_{t+1}^{*L|H}\right)}{u'\left(c_{t+1}^{*L|L}\right)}\right]^{-1} = \left[F_k\left(K_{t+1}^*,Z_{t+1}^*\right)+\left(1-\delta_k\right)\right], \qquad (3c)$$

$$v'\left(\frac{z_t^{*L}}{h_t^{*L|L}}\right)\frac{1}{h_t^{*L|L}}\left[1-\frac{\mu}{\pi^L}\frac{v'\left(\frac{z_t^{*L}}{h_t^{*L|H}}\right)\frac{1}{h_t^{*L|H}}}{v'\left(\frac{z_t^{*L}}{h_t^{*L|L}}\right)\frac{1}{h_t^{*L|L}}}\right]\frac{1}{u'\left(c_t^{*L|L}\right)}\left[1-\frac{\mu}{\pi^L}\frac{u'\left(c_t^{*L|H}\right)}{u'\left(c_t^{*L|L}\right)}\right]^{-1} = F_z\left(K_t^*,Z_t^*\right). \qquad (3d)$$

In (3c) and (3d), for low-type reporters, the social planner equalizes the MRS between consumption today and tomorrow to the MRT between consumption and investment today, and equalizes the MRS between consumption and leisure today to the MRT between consumption and labor today.

These conditions determine the constrained efficient allocation of the social planer.

3.2 Properties of the Optimum and Wedges

In the second best, marginal distortions in agents' choices may be understood using wedges. Agents choose effective labor and total spending, but the government observes neither work effort and consumption nor educational spending. Thus, there are two marginal distortions, defined as the intratemporal wedge (labor wedge) and the intertemporal wedge (capital wedge).

The labor wedge (intratemporal wedge)

$$\tau_{z_t}^i \equiv 1 - \frac{v'\left(\frac{z_t^{*i}}{h_t^{*i|i}}\right)\frac{1}{h_t^{*i|i}}}{u'\left(c_t^{*i|i}\right)}\frac{1}{F_z\left(K_t^*,Z_t^*\right)}. \qquad (4a)$$

The capital wedge (intertemporal wedge)

$$\tau_{k_{t+1}}^i \equiv 1 - \frac{u'\left(c_t^{*i|i}\right)}{\beta u'\left(c_{t+1}^{*i|i}\right)}\frac{1}{F_k\left(K_{t+1}^*,Z_{t+1}^*\right)+\left(1-\delta_k\right)}. \qquad (4b)$$

These wedges measure distortions of the second-best allocation relative to the laissez-faire allocation. In the dynamic taxation literature, the labor wedge or the labor distortion is the intratemporal wedge, which is defined as one minus the ratio of the MRS between consumption and labor today to the MRT between consumption and labor today. In a laissez-faire economy, the labor wedge is zero, since the agent would choose the allocation that equates the MRS today to the MRT today. However, as the spending in human capital and hence, the spending in consumption, is not observed by the government, the labor wedge may not be zero in the social planner's problem. While a positive labor wedge indicates that labor is distorted downwards, a negative labor wedge means that labor is distorted upwards. In a similar fashion,

the capital wedge is the intertemporal wedge, which is defined as the ratio of the MRS in consumption between today and tomorrow to the MRT between consumption today and investment today.

Substituting (3a) and (3b) into (4a) and (4b), it is clear that the intertemporal wedge and the intratemporal wedge both are zero for the high type: $\tau_{k_t}^H = \tau_{z_t}^H = 0$. This result confirms the familiar property of "no distortion at the top" in static models of Mirrlees (1971) and Stiglitz (1982), which states that the consumption-labor decision made by the truth-telling, highest-skill agent should be undistorted. Our study is valuable in that the result of "no distortion at the top" is robust in a dynamic setting, when the top type's skills is endogenously increased by investment, as opposed to being exogenously given in the existing literature. Intuitively, the allocation to a truth-telling, high type is not distorted by informational frictions, because low-type agents have no incentives to misreport as the high type. Thus, intertemporal and intratemporal wedges both are zero for an agent who reports as the high type.

By contrast, due to informational frictions, the social planner cannot distinguish truth-telling low types (i.e., $L|L$) from high types misreporting as low types (hereafter, misreporting high types) (i.e., $L|H$). This causes a non-zero wedge for an agent who reports as the low type. For these agents, we use (3c) and (3d) to rewrite the intertemporal and the intratemporal wedge in (4a) and (4b), respectively, as follows:

$$\tau_{k_{t+1}}^L = 1 - \frac{\left[1 - \dfrac{\mu}{\pi^L} \dfrac{u'\left(c_{t+1}^{*L|H}\right)}{u'\left(c_{t+1}^{*L|L}\right)}\right]}{\left[1 - \dfrac{\mu}{\pi^L} \dfrac{u'\left(c_t^{*L|H}\right)}{u'\left(c_t^{*L|L}\right)}\right]}, \quad \tau_{z_t}^L = 1 - \frac{\left[1 - \dfrac{\mu}{\pi^L} \dfrac{u'\left(c_t^{*L|H}\right)}{u'\left(c_t^{*L|L}\right)}\right]}{\left[1 - \dfrac{\mu}{\pi^L} \dfrac{v'\left(\frac{z_t^{*L}}{h_t^{*L|H}}\right)\frac{1}{h_t^{*L|H}}}{v'\left(\frac{z_t^{*L}}{h_t^{*L|L}}\right)\frac{1}{h_t^{*L|L}}}\right]}, \quad \text{for} \quad t = 1...T-1. \tag{5}$$

These wedges imply that $\tau_{k_{t+1}}^L > 0$ if and only if

$$MRS_{t,t+1}^{L|H} \equiv \frac{u'\left(c_t^{*L|H}\right)}{u'\left(c_{t+1}^{*L|H}\right)} < \frac{u'\left(c_t^{*L|L}\right)}{u'\left(c_{t+1}^{*L|L}\right)} \equiv MRS_{t,t+1}^{L|L}, \tag{6a}$$

and that $\tau_{z_t}^L \geq 0$ if and only if

$$MRS_{l,c}^{L|H} \equiv \frac{v'\left(\frac{z_t^{*L}}{h_t^{*L|H}}\right)\frac{1}{h_t^{*L|H}}}{u'\left(c_t^{*L|H}\right)} \leq \frac{v'\left(\frac{z_t^{*L}}{h_t^{*L|L}}\right)\frac{1}{h_t^{*L|L}}}{u'\left(c_t^{*L|L}\right)} \equiv MRS_{l,c}^{L|L}. \tag{6b}$$

Conditions (5) and (6a) suggest that, when the low type's *MRS* in consumption between today and tomorrow is greater than the corresponding *MRS* for the misreporting high type, the capital wedge is positive. Intuitively, a misreporting high type wants to saves more than the optimum, because the gain in consumption from reducing the expense on education decreases over time, but she has to restrict her saving as the same as the low type to avoid being detected by government. A positive capital tax on the

11

low type worsens her situation even more, hence it serves to correct the distortion, so as to encourage the high type to report the truth type.

Similarly, according to (5) and (6b), when a low type's *MRS* between labor and consumption today is greater than the corresponding *MRS* for a misreported high type, the labor wedge is positive. Intuitively, when a misreporting high type supplies labor less than the optimum, a positive labor tax on the low type serves to correct the distortion by inducing the low type to decrease the labor supply.

These capital and labor wedges for a low type characterize the smallest welfare cost that prevents a high type from misreporting as a low type. To determine the sign of those wedges, we need to analyze agents' behavior. In the next subsection, we study the optimal choice of the agent with a strategy $\sigma = (r|i)$.

3.3 Optimal Behavior of the Agent

Because agents' expenses on education are non-verifiable, agents with a strategy $\sigma = (r|i)$ can reallocate expenses between consumption and education without being caught by the social planner, as long as the sum of these two expenses is consistent to the announced type.

Given that the social planner distributes the constrained efficient allocation $\{c_t^{*r|r} + e_t^{*r|r}, k_t^*, z_t^{*r}\}_{t=1}^T$, an agent with a strategy $\sigma = (r|i)$ will choose the allocation $\{c_t^{*r|i}, e_t^{*r|i}, h_t^{*r|i}\}$ to maximize the problem:

$$\max_{\{c_t, e_t\}_{t=1}^T, \{h_t\}_{t=2}^T} \sum_{t=1}^T \beta^{t-1} \left[u(c_t) - v\left(\tfrac{z_t^{*r}}{h_t}\right) \right],$$

$$\text{s.t.} \quad c_t + e_t = c_t^{*r|r} + e_t^{*r|r} \quad \text{and} \quad h_{t+1}^i = \psi_t^i(h_t, e_t), \quad \text{with} \quad h_1^{r|i} = h_1 \quad \text{given.}$$

According to Definition 1, the constrained efficient allocation $\{c_t^{*r|i}, h_t^{*r|i}, e_t^{*r|i}\}_{t=1}^T$ must solve the above problem. The following theorem can be proved directly from the first-order conditions of the above problem.[9]

Theorem 1. The constrained efficient allocation A^* satisfies the following conditions:

for any $(r|i) \in \{(H|H), (L|L), (L|H)\}$,

1) for the terminal period $t=T$, $e_T^{*r|i} = 0$;

2) for the period before the terminal period $t=T-1$,

$$u'\left(c_{T-1}^{*r|i}\right) = \beta v'\left(\tfrac{z_T^{*r}}{h_T^{*r|i}}\right) \frac{z_T^{*r}}{\left(h_T^{*r|i}\right)^2} \frac{\partial}{\partial e}\left[\psi_{T-1}^i\left(h_{T-1}^{*r|i}, e_{T-1}^{*r|i}\right)\right]; \tag{7a}$$

3) for all other periods $t=1, 2, \dots, T-2$,

---

[9] All the proofs for the theorems and lemmas below are relegated in Appendix B.

$$u'\left(c_t^{*r|i}\right) = \left\{ \beta v'\left(\frac{z_{t+1}^{*r}}{h_{t+1}^{*r|i}}\right)\frac{z_{t+1}^{*r}}{\left(h_{t+1}^{*r|i}\right)^2} + \sum_{q=t+2}^{T}\left[\beta^{q-t} v'\left(\frac{z_q^{*r}}{h_q^{*r|i}}\right)\frac{z_q^{*r}}{\left(h_q^{*r|i}\right)^2}\prod_{s=t+1}^{q-1}\frac{\partial}{\partial h}\left[\psi_s^i\left(h_s^{*r|i},e_s^{*r|i}\right)\right]\right]\right\}\frac{\partial}{\partial e}\left[\psi_t^i\left(h_t^{*r|i},e_t^{*r|i}\right)\right]. \text{(7b)}$$

The theorem says that, in any period before the terminal period, *t*=1, 2 …*T*-1, an agent with strategy σ=(*r*|*i*) invests in human capital to the level, wherein the marginal cost of the investment in a particular period, which is the marginal utility of foregone consumption, is equal to the discounted sum of future marginal benefit of labor saving as resulted from an increase in human capital due to the investment. Hence, in period *T*-1, an agent invests in human capital to the level, wherein the foregone marginal utility of consumption equals the reduced discounted marginal disutility of labor in period *T*, evaluated at the increase in human capital. Of course, in the terminal period of the life cycle, no one invests to accumulate human capital, since there is no next period.

4. The Signs of the Wedge

In this section, we use a linear technology of human capital and show the signs of the wedges.[10] The following two assumptions are made.

Assumption 1. The disutility of work is iso-elastic: $v(l) = \kappa l^\gamma$, with $\kappa > 0$ and $\gamma > 1$.

In the existing Mirrlees models, an iso-elastic disutility of work has been widely used to prove that the labor wedge is positive and constant across periods and states, dubbed as "perfect labor tax smoothing" (e.g., Golosov et al., 2006; Werning, 2007; Farhi and Werning, 2013).

Assumption 2. The technology of human capital is of the form:

$$\psi_t^i\left(h_t^i,e_t^i\right) = \Phi\left(h_t^i,e_t^i\right) + \varepsilon_t^i \quad \text{for } i=H, L, \text{ with } \varepsilon_1^H > \varepsilon_1^L = 0 \quad \text{and} \quad \varepsilon_t^H = \varepsilon_t^L = 0 \quad \text{for } t \geq 2, \tag{8}$$

where $\Phi\left(h_t^i,e_t^i\right) = \left(1-\delta_h\right)h_t^i + \alpha e_t^i$, with $0 \leq \delta_h < 1$ and $\alpha > 0$.

The technology of human capital is linear.[11] Notice that the technology is the same for different

---

[10] In Section 6, we numerically show that, even if the technology of human capital is non-linear, the results obtained in this section also hold.

[11] Although this linear form gives $\frac{\partial^2}{\partial h \partial e}\psi_t^i(h,e) = 0$, we should also mention that the existing literature has not come to an agreement concerning the sign of $\frac{\partial^2}{\partial h \partial e}\psi_t^i(h,e)$. For example, Kapicka (2006) used the Cobb-Douglas form which indicates that $\frac{\partial^2}{\partial h \partial e}\psi_t^i(h,e)$ is positive, while da Costa and Maestri (2007) stated that $\frac{\partial^2}{\partial h \partial e}\psi_t^i(h,e)$ should negative if education decreases wage variance. To carry out a robustness check, in Appendix A we offer an analysis in a general function form in a two-period model. Moreover, in Section 6, we provide a numerical analysis using the

skill types except initial $\varepsilon_1{}^H > \varepsilon_1{}^L$. The initial learning difference is the only heterogeneity among agents. Two remarks are in order.

First, a linear technology is the simplest functional form that captures the law of motion for human capital and also renders a tractable framework. There is a vast literature on human capital accumulation that uses a linear technology to analyze and estimate the models. See, for example, Todd and Wolpin (2003, 2007), Cunha and Heckman (2008) and Vogl (2016).

Second, Todd and Wolpin (2003, 2007) and Cunha and Heckman (2008) take the term $\varepsilon_1{}^i$ as some unobservable effects at stages of a child's life cycle. The difference captures some innate abilities and cognitive and noncognitive skills in early childhood development. In our model, instead of taking on agents with idiosyncratic uncertainty on skills across time, we posit that agents are born with identical human capital except initial learning abilities which endogenously generate heterogeneities on skills over time. As these initial disparities fade over time, skill differences are endogenously determined over time. To illustrate how initial disparities fade over time, we follow Cunha and Heckman (2007) and rewrite (8) in a recursive form as follows.

$$h_{t+1}^i = \alpha e_t^i + \alpha(1-\delta_h)e_{t-1}^i + ... + \alpha(1-\delta_h)^{t-1}e_1^i + (1-\delta_h)^t h_1^i + (1-\delta_h)^t \varepsilon_1^i.$$

It is clear that the weight of the initial disparity associated with the term $\varepsilon_1{}^i$ declines over time.

We are ready to analyze the sign of the wedge for agents who announce to be low-skilled. We start by establishing three lemmas. Since agents have no incentives to invest in human capital in the terminal period, truth-telling low types (i.e., $L|L$) and misreporting high types (i.e., $L|H$) all allocate the same amount of income in consumption. Consumption is observable in the terminal period.

Before the terminal period, agents invest in human capital. Due to the fact that expenses on education are not observable, those misreporting high types may invest less in human capital and spends more on consumption. The following lemma confirms that in period T-1, a misreporting high type invests less in human capital and spends more on consumption than the low type.

Lemma 1. Under Assumptions 1 and 2, the constrained efficient allocation A^* satisfies $u'(c_{T-1}^{*L|H}) < u'(c_{T-1}^{*L|L})$.

Building on the result of Lemma 1, we can show that, for periods t=1, 2… T-1, misreporting high types indeed invest in human capital less than the low type, but their levels of human capital (or skills) are still greater than the low type. Moreover, the human capital of high-type deviators and that of the

---

Cobb-Douglas form. All the results are consistent with the results obtained in this section.

low type will get closer over time, but they never cross each other. We state the result as follows.

Lemma 2. Under Assumptions 1 and 2, the constrained efficient allocation A^* satisfies $e_t^{*L|H} < e_t^{*L|L}$ and $h_{t+1}^{*L|L} < h_{t+1}^{*L|H}$ for t=1 ... T-1. Furthermore,

$$\frac{h_2^{*L|L}}{h_2^{*L|H}} < \frac{h_3^{*L|L}}{h_3^{*L|H}} < ... < \frac{h_{T-1}^{*L|L}}{h_{T-1}^{*L|H}} < \frac{h_T^{*L|L}}{h_T^{*L|H}} < 1.$$

Finally, based on Lemma 2, we verify that a misreporting high type allocates more expenses on consumption than the low type in all except the last period, and the ratio of their marginal utility of consumption increases over time.

Lemma 3. Under Assumptions 1 and 2, the constrained efficient allocation A^* satisfies

$$\frac{u'\left(c_1^{*L|H}\right)}{u'\left(c_1^{*L|L}\right)} < \frac{u'\left(c_2^{*L|H}\right)}{u'\left(c_2^{*L|L}\right)} < ... < \frac{u'\left(c_{T-1}^{*L|H}\right)}{u'\left(c_{T-1}^{*L|L}\right)} < \frac{u'\left(c_T^{*L|H}\right)}{u'\left(c_T^{*L|L}\right)} = 1.$$

Lemma 3 indicates that the marginal rate of substitution in consumption between two consecutive periods for low types is larger than that for misreporting high types, i.e., $MRS_{t,t+1}^{L|L} > MRS_{t,t+1}^{L|H}$ for t=1… T-1. It is clear from (3c) and (6a) that the inverse Euler equation does not hold and thus, the intertemporal wedge must be positive for the low type in t=1… T-1.

We are ready to analyze the wedges. First, we establish the proposition for the intertemporal wedge.

Proposition 1. Suppose that Assumptions 1 and 2 hold. Then, the inverse Euler equation does not hold for the low type and has the following relation

$$\frac{1}{u'\left(c_t^{*L|L}\right)} > \frac{1}{\beta\left[F_k\left(K_{t+1}^*, Z_{t+1}^*\right)+\left(1-\delta_k\right)\right]u'\left(c_{t+1}^{*L|L}\right)}, \quad \text{for } t=1… T\text{-}1.$$

Proposition 1 implies that the intertemporal wedge for the low type is always positive ($\tau_{k_{t+1}}^L > 0$). To understand the reason, as the spending in human capital is non-verifiable, misreporting high types not only reduce labor supply for leisure, but they also decrease spending in human capital for consumption. The reduction in human capital investment will gradually make their skills close to those of the low type. With the initial advantage fading over time, misreporting high types cannot lower human capital investment as much as before. Diminishing skills in the future make misreporting high types have urge to save more. To discourage high types from deviation so as to push them to invest in human capital, the capital wedge for the low type is positive, because it cancels out the benefit from deviation rapidly. Thus,

the taxation on physical capital for the low type is an efficient way to correct the incentives of the high type and fosters human capital.

This result is in a sharp contrast to the zero intertemporal (capital) wedge in the existing Mirrlees models without human capital (e.g., Goloslov et al. 2006; Werning, 2007) and with human capital (e.g., Da Costa and Maestri, 2007; Anderberg, 2009; Stantcheva, 2016). In these models, if there are no uncertainties on skills, misreporting high types cannot reallocate more resources to consumption, as consumption can be inferred from savings and labor income. Then, high-type deviators have the same consumption in each period and thus consumption is smooth over time; so do low types. Hence, high-type deviators and low types have the same marginal rate of substitution in consumption, which gives a zero intertemporal wedge.

Next, we state the proposition that establishes the sign of the intratemporal wedge for the low type.

Proposition 2. Suppose that Assumptions 1 and 2 hold. Then, the intratemporal wedge for the low type is negative in the first period and is positive in the terminal period; that is, $\tau^L_{k_{t+1}} > 0$ and $\tau^L_{z_T} > 0$.

Intuitively, in the first period, although both types have the same level of human capital, high types have an advantage to accumulate human capital in the first period. Thus, high-type deviators need not invest in human capital as much as low types. The reduction in human capital investment would render high-type deviators to consume more. A negative labor wedge on low types induce low types to work more. Such a policy encourages high types to tell the truth, because if not, by cutting spending in human capital investment, high-type deviators increase the level of consumption that is already high which only increases the utility by a small margin. Yet, by mimicking low types, extra labor would decrease their utility by a large margin. As a result, a negative labor wedge on low types in the first period serves as a mechanism to induce high types to invest more in human capital.

Conversely, in the terminal period, both types have no incentives to invest in human capital. This goes back the scenario of the Mirrlees literature. If high types misreport as low types, they could have more leisure and less labor than low types. Then, high-type deviators and low types have different marginal utilities of leisure. To punish high types who misreport as low types, the social planner levies labor income taxes on low types, which may discourage low types from working. Yet, the policy hurts high-type deviators more than low types and thus, is an efficient way to induce high types to tell the truth and work more.

Between the first and the terminal periods (i.e., $t=2\ldots T$-1), the intratemporal wedge is ambiguous because, in these periods, high-skilled deviators would gain from reducing both human capital investment

16

and the labor supply. If the former effect dominates, a negative labor wedge would discourage high-skilled agents from deviation and is optimal. By contrast, if the latter effect dominates, a positive labor wedge is optimal. We are not sure of which effect dominates.

This is a new result in the dynamic Mirrlees literature. In the existing literature, in order to prevent the high-skilled from misreporting their type, the labor wedge for the low type is positive and constant in each period (e.g., Golosov et al., 2006; Werning, 2007; Farhi and Werning, 2013). Our model does not have the perfect labor tax smoothing. In particular, the labor wedge for the low type is negative in the first period and ambiguous in all other periods before the terminal period. These different results emerge, because the human capital investment is endogenous and non-verifiable. A negative labor wedge on the low type serves as a mechanism that prevents high types from misreporting as low types and induce them to invest in human capital. We note that in models that considered an extensive labor supply, Diamond (1980) and Saez (2002) also obtained a negative labor income tax at the bottom income. In these papers, a subsidy for the working poor is optimal, because the participation effect of the labor force dominates the incentive effect of higher income earners, which is different from the effect on human capital investment in our model.

5.   Implementation

Although it is tempting to interpret the wedges defined in (4a) and (4b) as actual taxes on labor and capital, the relationship between wedges and taxes is not straightforward. Each wedge controls only one aspect of agent's behavior (labor or savings in a particular period), taking all other choices fixed *at the optimal level*. However, agents choose all aspects (labor and savings) jointly. In this section, we propose a tax system that implements the constrained efficient allocation as a competitive equilibrium.

There are different tax systems that can implement the constrained efficient allocation as a competitive equilibrium. Golosov et al. (2006) pointed out that the simplest way of implementation is to assign arbitrarily high punishments if agents' observable allocation in any period is different from the constrained efficient allocation. Yet, this way severely limits an agent's choices and may be unrealistic. To relax the limitation and provide a direct connection between wedges and optimal tax rates, this paper uses a history-dependent tax system that punishes agents if their history of capital k^t and effective labor z^t does not satisfy some conditions.[12] If these conditions are met, in addition to lump-sum taxes, our tax system uses linear tax rates to levy labor income and capital income. We show that these optimal linear tax rates are exactly the same as the wedges that are established by the social planner and derived in

---

[12]   Below, we denote by x^t the history of x up to period t; e.g., $x^t \equiv (x_1, x_2, .., x_t)$.

subsection 3.2.

5.1 Agent's Problem

Given the price $\{r_t, w_t\}_{t=1}^T$, tax system $\{T_t^e(k^t, z^t)\}_{t=1}^T$, and initial physical and human capital (k_1, h_1), the problem of an agent of type $i \in \{H, L\}$ is to maximize

$$\max \sum_{t=1}^T \beta^{t-1} \left[u(c_t) - v(l_t) \right],$$

subject to budget constraints and the human capital accumulation as follows:

$$c_t + e_t + k_{t+1} \le w_t h_t l_t + (1+r_t)k_t - T_t^e(k^t, z^t), \quad t=1, 2, \ldots, T, \quad k_{T+1} = 0$$

$$h_t = \psi_{t-1}^i(h_{t-1}, e_{t-1}), \quad t=2, 3, \ldots, T.$$

5.2 Firm's Problem

Given the price $\{r_t, w_t\}_{t=1}^T$, the representative firm solves the following problem:

$$\max_{K_t, Z_t} F(K_t, Z_t) - w_t Z_t - (r_t + \delta_k)K_t \quad \text{for } t=1, 2, \ldots, T.$$

The first-order conditions are

$$w_t = F_z(K_t, Z_t),$$

$$r_t = F_k(K_t, Z_t) - \delta_k.$$

5.3 The Government

The government balances the budget as follows.

$$\sum_{i=H,L} \pi^i T_t^e(k^{it}, z^{it}) \ge G_t, \quad \text{for } t=1, 2, \ldots, T,$$

where k^t and z^t are the history of capital and effective labor allocation for a type i agent.

5.4 Competitive Equilibrium

We define a competitive equilibrium and an optimal tax system. Given a tax system $\{T_t^e(k^t, z^t)\}_{t=1}^T$ and government expenditure $\{G_t\}_{t=1}^T$, a **competitive equilibrium** is an allocation $A^e = (c^e, e^e, h^e, l^e, z^e, k^e, Z^e, K^e)$ and prices $\{r_t, w_t\}_{t=1}^T$ such that, (1) given prices $\{r_t, w_t\}_{t=1}^T$, $\{c_t^{ie}, e_t^{ie}, h_t^{ie}, l_t^{ie}, z_t^{ie}, k_t^{ie}\}_{t=1}^T$ solves the problem of an agent of type $i=H, L$; (2) prices $\{r_t, w_t\}_{t=1}^T$ are given by $r_t = F_k(K_t, Z_t) - \delta_k$ and $w_t = F_z(K_t, Z_t)$ for $t=1, 2, \ldots, T$; and (3) markets clear and thus, A^e is a feasible allocation.

18

Definition 2. A tax system $\{T_t^e(k^t, z^t)\}_{t=1}^T$ is **optimal** if it implements the constrained efficient allocation A^* as a competitive equilibrium allocation A^e.

We denote $\{T_t^{e*}(k^t, z^t)\}_{t=1}^T$ as the optimal taxation.

5.5 Implementation with a tax system

We now propose a tax system that implements the constrained efficient allocation A^* in a competitive equilibrium. The tax system includes linear income tax rates $(\tilde{\tau}_{z_t}^L, \tilde{\tau}_{k_t}^L)$ and lump-sum taxes Γ_t^L if condition $\rho_t^L(k_t - k_t^{*L}) = z_t^{*L} - z_t$ is met, and linear income tax rates $(\tilde{\tau}_{z_t}^H, \tilde{\tau}_{k_t}^H)$ and lump-sum taxes Γ_t^H if condition $\rho_t^H(k_t - k_t^{*H}) = z_t^{*H} - z_t$ is met. If none of these two conditions are met, then the agent will be punished severely enough.

The tax system is:

$$
T_t^{e*}\left(k^t, z^t\right) = \begin{cases} \tilde{\tau}_{k_t}^L\left(1+r_t\right)k_t + \tilde{\tau}_{z_t}^L w_t z_t + \Gamma_t^L & \text{if } \rho_s^L\left(k_s - k_s^{*L}\right) = z_s^{*L} - z_s \text{ for all } s \leq t \\ \tilde{\tau}_{k_t}^H\left(1+r_t\right)k_t + \tilde{\tau}_{z_t}^H w_t z_t + \Gamma_t^H & \text{if } \rho_s^H\left(k_s - k_s^{*H}\right) = z_s^{*H} - z_s \text{ for all } s \leq t \\ \infty & \text{otherwise} \end{cases}
$$

These two conditions are explained as follow. First, linear tax rates $(\tilde{\tau}_{z_t}^L, \tilde{\tau}_{k_t}^L)$ are created for low types, if their allocation (z^T, k^T) satisfies $\rho_t^L(k_t - k_t^{*L}) = z_t^{*L} - z_t$ for $t \leq T$. Second, linear tax rates $(\tilde{\tau}_{z_t}^H, \tilde{\tau}_{k_t}^H)$ are designed for high types, if their allocation satisfies $z_t = z_t^{*H}$ and any k_t for $t \leq T$.

The following theorem shows that this tax system implements the constrained efficient allocation as a competitive equilibrium and the linear tax rates on the equilibrium path are consistent with the wedges in Subsection 3.2.

Theorem 2 There exists an optimal tax system $\{T_t^{e*}(k^t, z^t)\}_{t=1}^T$ such that the linear income tax rates are consistent with the wedges. That is,

$$
\tilde{\tau}_{k_t}^i = 1 - \frac{u'\left(c_{t-1}^{*i|i}\right)}{\beta\left(1+r_t^*\right)u'\left(c_t^{*i|i}\right)} \quad \text{for } t = 2,3,...,T, \text{ and } i = H,L, \tag{9a}
$$

$$
\tilde{\tau}_{z_t}^i = 1 - \frac{v'\left(\frac{z_t^{*i}}{h_t^{*i|i}}\right)\frac{1}{h_t^{*i|i}}}{w_t^* u'\left(c_t^{*i|i}\right)} \quad \text{for } t = 1,2,...,T, \text{ and } i = H,L, \tag{9b}
$$

where $w_t^* = F_z\left(K_t^*, Z_t^*\right)$ and $r_t^* = F_k\left(K_t^*, Z_t^*\right) - \delta_k$.

6. Numerical Analysis

In this section, we offer numerical analysis. Our numerical analysis takes a middle position between a simple demonstration of the optimal mechanism and a careful calibration of quantitative implications for the optimal wedge.

For the technology of human capital, we use the Cobb-Douglas function taken from Ben-Porath (1967) which is more general than the linear function in Section 4. As we will see, the results in Propositions 1 and 2 are quantitatively robust, when the general form is used. Our quantitative exercises also shed light on new insights.

6.1 Calibration

We calibrate a baseline economy based on the US data and then quantitatively solve the constrained efficient allocation. The calibration proceeds as follows. We construct the baseline economy, which is a decentralized economy with two types of agents and with linear tax rates on capital and labor income. Agents have three periods of lives and each period represents 15 years.[13] For simplicity, the fraction of each type is assumed to be equal, i.e., $\pi^H=\pi^L=0.5$. In the baseline economy, some parameter values are set based on the exiting literature, normalization or assumptions. Table 1 lists these parameter values. The rest of the parameter values are determined endogenously to match with the data. Table 2 summarizes these endogenously calibrated parameter values. We apply these parameter values to the second-best economy and calculate the constrained efficient allocation, capital wedges and labor wedges. In addition, we compare the capital and labor wedge between models with and without endogenous human capital evolution.

For the tax system of the baseline economy, according to McDaniel (2007), the average tax rates in the US during 1960-2007 for the capital income and the labor income are around 0.3 and 0.2, respectively. Thus, in our baseline economy, we set $\tau_k=30\%$ and $\tau_z=20\%$. We assume zero government expenditure $G_t=0$ in every period, so the tax revenue is equally redistributes to the two types of agents as a lump-sum transfer LS_t. The budget constraint for an agent is as follows:

$$c_t + e_t + k_{t+1} = \left\{1+\left(1-\tau_K\right)\left[F_K\left(K_t,Z_t\right)-\delta_k\right]\right\}k_t + \left(1-\tau_Z\right)F_Z\left(K_t,Z_t\right)z_t + LS_t, \ t=1,2,3,$$

where $e_3=0$ and $k_4=0$, because the third period is the terminal period.

The periodic utility function is of the form:

---

[13] We assume that agents enter the economy at age 20 and retire at age 66. The first period represents ages between 20 and 35, the second period ages between 36 and 50 and the third period ages between 51 and 65.

$$u\left(c_t\right) - v\left(l_t\right) = \frac{c_t^{1-\sigma}}{1-\sigma} - \kappa\left(\frac{z_t}{h_t}\right)^{\gamma}.$$

Following Farhi and Werning (2013), we set $\gamma=3$ and $\kappa=1/3$, which implies the Frisch elasticity for labor of 0.5. Following Conesa et al. (2009), we set $\sigma=2$, implying the intertemporal elasticity of substitution of 0.5. The discount factor between periods is set to be $\beta=(0.96)^{15}=0.542$. Also, the initial human capital level is normalized to be $h_1=1$, meaning that the effective labor supply is equal to labor hours for both types in the first period.

The technology of the final good is assumed to be the Cobb-Douglas form:

$$Y = F\left(K,Z\right) = AK^{\xi}Z^{1-\xi}. \tag{10}$$

Following Conesa et al (2009), we normalize $A=1$ and set $\xi=0.36$. These authors set the annual depreciation rate to be 0.0833. We go along with the value and translate it to $\delta_k=0.729$ for 15 years.

Recall that the technology of human capital takes the form $\psi_t^i(h_t^i, e_t^i) = \Phi(h_t^i, e_t^i) + \varepsilon_t^i$, $h_1^H = h_1^L$, $\varepsilon_1^H > \varepsilon_1^L$ and $\varepsilon_t^H = \varepsilon_t^L = 0$ for $t \geq 2$. In section 4, we use a linear form for $\Phi(h_t^i, e_t^i)$ in order to obtain analytical solution. To offer quantitative analysis, here we adopt the Cobb-Douglas form taken from Ben-Porath (1967):

$$\Phi\left(h_t^i, e_t^i\right) = \left(1-\delta_h\right)h_t^i + \alpha(h_t^i)^{\eta_1}(e_t^i)^{\eta_2}. \tag{11}$$

The form reduces to a linear form if $\eta_1=0$ and $\eta_2=1$. For parameter values, we normalize $\alpha=1$. Following Manuelli and Seshadri (2014), we set $\eta_1=0.486$, $\eta_2=0.4$ and $\delta_h=0.337$. Without loss of generality, we set $\varepsilon_1^L=0$, which simplifies the notion that the high type has an initial learning advantage.

The initial aggregate physical capital K_1 and the initial advantage of human capital ε_1^H for the high type are calibrated. We calibrate these two parameters to match the data of the capital-output ratio and the wage premium, respectively. Following Peterman (2016), annul capital-output ratio in the US is 2.7, which implies a 15-year capital-output ratio of 0.18. The estimated value for the wage premium in the literature lies within 1.2 and 2.4, such as 1.26-1.74 in Murphy and Welch (1992), 1.37-1.75 in Autor et al. (1998), 1.7-2.4 in Heathcote et al. (2010), and 1.2-2.2 in James (2012). Existing research indicates that the wage premium has risen substantially in the recent years. To uncover the optimal capital income tax and labor income tax in response to the rising wage premium, our calibration targets different values within the estimated range of the wage premium: the lowest value of 1.2, a medium value of 1.8 and the highest value of 2.4.

As our model does not have a steady state, we cannot calibrate our model in the same way as the model with a steady state. We use the simulation method to calibrate the values of K_1 and ε_1^H in the way

as follows. Using the parameter values in Table 1, we simulate the baseline economy along with the use of the two technologies in (10) and (11). The optimal allocation targets the capital-output ratio at a value around 0.18 and the wage premium in period 3 at one of the three values, 1.2, 1.8 and 2.4. The resulting calibrated value for K_1 is 0.081 and the resulting calibrated values for ε_1^H are 0.126, 0.463 and 0.756. (See Table 2).[14] Given these different calibrated values, in the numerical analysis below we solve the constrained social problem in (1a)-(1c) and then compute the wedge and the optimal constrained allocation in the range of $\varepsilon_1^H \in [0, 0.8]$.

6.2 Results

Endogenous vs exogenous human capital: Now, we analyze the wedges and compare with the model with exogenous human capital.

First, in our model with endogenous human capital, we compute the capital wedge and the labor wedge in the range of $\varepsilon_1^H \in [0, 0.8]$. As expected, both wedges for high types are zero. For low types, both the capital and labor wedges are reported as solid lines in Figures 1 and 2, respectively. It is clear that capital wedges for low types are positive (cf. Figure 1)[15]. Moreover, the labor wedge for low types is negative in the first period and positive in the terminal period (cf. Figure 2). Thus, the results summarized in Propositions 1 and 2 based on a linear human capital technology are quantitatively robust to the Cobb-Douglas technology.

We note that the capital wedge and the negative labor wedge are relatively small, because the distortion is caused by the unobservable human capital investment. As in Stantcheva (2016), this is akin to the "production efficiency" result in that human capital investment is an intertemporal decision with persistent effects. Hence, it is relatively costly to distort the decision for human capital investment for redistributive reasons, unless the redistributive effect is very strong.

[Insert Figures 1 and 2 here.]

Next, to compare with the wedge in the model with exogenous human capital, we shut down the decision for human capital investment. Specifically, the evolution of human capital for the two types of agents are treated otherwise the same as the counterpart in our model, except no decision made for

---

[14] To rule out optimal allocations of complex numbers in the problem, we use the program *fmincon* in the Matlab. The program solves a minimization problem with linear or nonlinear constraints and restricts solutions to real numbers. Substituting the optimal solutions derived by the program to all the necessary conditions, we find that the error is less than an order of 10^{-6}.

[15] Since the initial physical capital K_1 is predetermined, the capital wedge in the first period plays no rules in the model. Thus we only present capital wedges after the first period.

human capital investment. In such a model, we find that capital wedges for low types are zero, as plotted in dashed lines in Figure 1. Intuitively, a zero capital wedge for low types arises, because capital is not distorted in the model with exogenous human capital. This is reminiscent of the result of a zero capital wedge in the existing dynamic Mirrlees model when there are no idiosyncratic shocks (e.g., Golosov et al., 2006: Stantcheva, 2016). The labor wedge is also zero in the first period (cf. dashed lines in Figure 2). Because we set both types of agents to have the same initial human capital, the labor wedge is not distorted in the first period. Nevertheless, the labor wedge is positive in later periods. Like the existing dynamic Mirrlees models without human capital (e.g., Golosov et al., 2006), as high types have higher skills in later periods, to avoid high types to mimic low types and to have more leisure, positive labor wedges are optimal in all these periods. Note from the figure that these positive labor wages are larger than those in the model with endogenous human capital. Intuitively, with endogenous human capital investment, the optimal marginal tax rates on labor income are smaller. As the low type also chooses human capital investment, a smaller labor wage serves to encourage low types' incentives in human capital investment. In conclusion, we find that capital wedges are higher and labor wedges are lower when human capital is endogenous.

Moreover, with endogenous human capital, the capital wedges feature a bell shape in high types' initial advantages in accumulating skills, ε_1^H: they increase first and then decrease (cf. Figure 1). Intuitively, when high types' initial learning advantages are increasing, high types can reduce more expenses on education in order to increase more expenses on consumption and thus, they have greater incentives to mimic low types. An increasing capital wedge is necessary to discourage such incentives. However, the expenses on consumption and education distributed to the low type also gets smaller as high types' initial learning advantages are increasing, which reduces the benefit from mimicking as the low type. See Figures 3 and 4. As high types' initial learning advantages are larger than a threshold, the latter effect dominates, and thus the benefit from misreporting is smaller. As a result, a smaller capital wedge is sufficient to deter the high type's incentive to deviate.

[Insert Figures 3 and 4 here.]

Constrained efficient allocation: Figures 3 and 4 present constrained efficient consumption and human capital investment. According to Figure 3, the gap in consumption between high-type deviators and low types (c.f., $c_1^{L|H}$ and $c_1^{L|L}$) is increasing first then decreasing in high types' initial learning advantages. According to Figure 4, the gap in human capital investment between high-type deviators and low types (c.f., $e_1^{L|H}$ and $e_1^{L|L}$) which is also increasing first then decreasing in high types' initial learning advantages. With small initial advantages for the high type, when initial advantages increase, these gaps

in human capital investment get larger and high types have larger incentives to deviate. As such, the optimal capital wedge is increasing. When high types' initial advantages are greater than a threshold, these gaps get smaller and thus the incentive to deviate is smaller. Then, the optimal capital wedge is decreased.

Figure 5 presents constrained efficient effective labor. Of a particular note is that the low type provides more effective labor than the high type in the early stage. Intuitively, since both types have the same initial human capital, this indicates larger working effort of the low type in the early stage. This result is consistent with the observation that many low-skilled labors drop out schools to take full-time jobs earlier than those high-skilled labors. Thus, in our model, the labor wedge is negative in the first period, implying that low types' labor income should be subsidized, in order to prevent high types from reducing human capital investment (cf. Figure 2).

[Insert Figure 5 here.]

Figure 6 illustrates constrained efficient human capital level. In particular, the deviating high type is shown in the dotted line. The result indicates that, despite cutting expenses on education, because of initial learning advantages, the deviating high type's level of human capital lies between the truth-telling high type and the low type, which confirms Lemma 2.

[Insert Figure 6 here.]

Government expenditure: So far, our numerical exercises are carried out under the environment with zero government expenditure, G_t=0. We now quantify how the government expenditure affects the optimal taxation. Given that more resources are devoted to the government expenditure which is a waste in our setup, the government expenditure may be expected to increase the wedge. Golosov et al. (2006) has numerically shown that the capital wedge and the labor wedge are both increasing in the government expenditure. However, because human capital is endogenous in our model, this is not necessarily the case in the labor wedge. To see this, we set G_t=0.085 in each period.[16] The resulting labor wedge is in Figure 7, wherein the solid line is for G=0 and the dashed line is for G=0.085. We find that the labor wedge is decreasing, rather than increasing, in the government expenditure in the first period.

[Insert Figure 7 here.]

The reason goes as follows. Since high types have initial advantages to accumulate human capital, they may deviate and increase consumption by reducing expenses on education. As explained earlier, a negative labor wedge is optimal in the first period, which encourages low types to work more so as to prevent high types from mimicking low types. Now, when the government expenditure increases, average

---

[16] According to Peterman (2016), the government expenditure-output ratio is 17%. In our model, the output is around 0.5. Therefore, we target the government expenditure at 0.085.

consumption is lower and average labor is higher, which would increase high types' incentives to mimic low types in an attempt to work less. When the government subsidizes low types' labor more, this effectively prevents high types from misreporting as low types.

Welfare: Finally, we calculate the welfare of the constrained efficient allocation. In particular, we compare the welfare gain of the second-best economy from the baseline economy with linear income tax rates. Let the welfare in the baseline economy (*BE*) be denoted by $W^{BE}(\overline{c}_t^H, \overline{c}_t^L, \overline{l}_t^H, \overline{l}_t^L)$, where $\overline{c}_t^i$ and $\overline{l}_t^i$ are, respectively, consumption and the labor supply of type *i*=*H*, *L* in the baseline economy in time *t*. Let the welfare of the second-best economy (*SE*) be denoted by W^{SE}.

The welfare gain of the second-best economy from the baseline economy is defined in terms of consumption equivalence: the percentage increase in consumption in the second-best economy relative to the baseline economy. Denote ω as the percentage increase in consumption which satisfies:

$$W^{BE}\left((1+\omega)\overline{c}_t^H, (1+\omega)\overline{c}_t^L, \overline{l}_t^H, \overline{l}_t^L\right) = W^{SE}.$$

First, we compute the welfare gain of the second-best economy from the baseline economy, when high types' initial advantages in human capital accumulation ε_1^H are small, medium and large, represented by 0.126, 0.463 and 0.756, respectively. The results are reported in Table 3, where the welfare gain of the baseline economy from itself is zero in the first row. The second row is the welfare gain of the second-best economy from the baseline. The welfare gain increases in ε_1^H. Note that a larger ε_1^H corresponds to a larger wage premium and thus, a larger skill difference between high skills and low skills. The results suggest that the constrained efficient allocation gives a higher welfare in an economy wherein the skill difference is larger.

[Insert Table 3 here.]

Next, in our second-best economy, human capital investment and agents' types both are private information. It is interesting to see how the welfare gain over the baseline economy changes if the private information becomes public. We compute two otherwise the same second-best economies except for relaxing some informational distortions. In Table 3, the third row is the same economy except for observable expenses on human capital investment (HCI observable), and the fourth row is the same economy except for no private information, wherein human capital investment and agents' types both are observable (HCI and type observable). The results in Table 3 indicate that, although the welfare increases when the distortion is reduced, the gain is large when the skill types are observable and small when expenses on human capital investment are observable.

7.   Concluding Remarks

In the existing dynamic Mirrlees models with heterogeneous and unobservable skills which evolve exogenously over time, the government chooses the constrained efficient allocation which implies conditions that give rise to capital and labor wedges. Along the literature, a positive capital wedge is designed for the insurance purpose, and the wedge is zero if there is no idiosyncratic shock to skills and thus no insurance motive. Moreover, the labor wedge for low-skill types is positive and constant over time, in order to prevent high-skill types from misreporting as low-skilled types and smooth consumption over time.

In this paper, we revisit the optimal income tax policy in a dynamic Mirrlees model with heterogeneous skills. The key departure from the existing studies is that individuals' unobservable skills evolve over time by means of endogenous human capital investment, rather than through exogenous shocks. We design a tax system that can foster human capital investment. Since our model does not introduce any kinds of uncertainty, there is no insurance motive. Even so, our model yields a positive capital wedge. Moreover, the labor wedge is neither always positive nor constant over time. The labor wedge is negative in the first period and ambiguous in all other periods except the terminal period.

These wedges emerge from the setting of endogenous human capital investment that is not observable by the government. Positive capital wedges are optimal in our model, because the policy discourages high-type agents from mimicking low types so as to not consume too much as a result of reducing unobservable expenses on human capital investment. Negative labor wedges are optimal in early periods of agents' life cycle, because the policy attracts low types to work more, which deters high-type agents from mimicking low types so as to not work more and not consume more by way of reducing unobservable expenses on human capital investment.

We propose a history-dependent tax system to implement the wedges. In the tax system, capital and labor income are taxed linearly, along with lump-sum taxes, if an agent's history of capital and effective labor satisfies some conditions; otherwise, an agent would be taxed severely. Relative to the laissez-fare economy with linear taxes, our second-best optimal mechanism gives rise to a welfare gain, with a larger gain when the gap in agent's abilities increases.

References

Albanesi, S. and C. Sleet (2006) Dynamic optimal taxation with private information. *Review of Economic Studies* 73, 1-30.

Anderberg, D. (2009) Optimal policy and the risk properties of human capital reconsidered. *Journal of Public Economics* 93, 1017-1026.

Autor, D., L.F. Katz and A.B. Krueger (1998) Computing inequality: Have computers changed the labor market? *Quarterly Journal of Economics* 113(4), 1169-1213.

Becker, G.S. (1964), Human Capital: *A Theoretical And Empirical Analysis, With Special Reference To Education.* University of Chicago Press.

Ben-Porath, Y. (1967) The production of human capital and the life cycle of earnings. *Journal of Political Economy* 75(4), 352–365.

Boháček, R. and M. Kapička (2008) Optimal human capital policies. *Journal of Monetary Economics* 55, 1-16.

Bovenberg, A.L., and B. Jacobs (2005) Redistribution and education subsidies are Siamese twins. *Journal of Public Economics* 89, 2005-2035.

Conesa, J.C., S. Kitao, and D. Krueger (2009) Taxing capital? Not a bad idea after all! *American Economic Review* 99(1), 25–38.

Cunha, F. and J.J. Heckman (2007) The technology of skill formation. *American Economic Review* 97(2), 31–47.

Cunha, F. and J.J. Heckman (2008) Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources* 43, 738-782.

Da Costa, C. and L.J. Masestri (2007) The risk properties of human capital and the design of government policies. *European Economic Review* 51, 695-713.

Diamond, P. (1980) Income taxation with fixed hours of work. *Journal of Public Economics* 13, 101–110.

Diamond, P. A. (1998) Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review* 88(1): 83-95.

Ebert, U. (1992) A reexamination of the optimal nonlinear income tax, *Journal of Public Economics* 49, 47-73.

Farhi, E. and I. Werning (2013) Insurance and taxation over the life cycle. *Review of Economic Studies* 80, 596-635.

Farhi, E., C. Sleet, I. Werning and S. Yeltekin (2012) Non-linear capital taxation without commitment. *Review of Economic Studies* 79, 1469-1493.

Goldin, C.D. and L.F. Katz (2008) *The Race Between Education And Technology.* Harvard University Press.

Golosov, M., N.R. Kocherlakota and A. Tsyvinski (2003) Optimal indirect and capital taxation, *Review of Economic Studies* 70, 569-587.

Golosov, M., A. Tsyvinski, and I. Werning (2006) New dynamic public finance: A user's guide. *NBER Macroeconomic Annual*, MIT Press.

Golosov, M. and A. Tsyvinski (2006) Designing optimal disability insurance: A case for asset testing. *Journal of Political Economy* 114(2), 257-279.

Grochulski, B. and T. Piskorski (2010) Risky human capital and deferred capital income taxation. *Journal of Economic Theory* 145, 908-943.

Heathcote, J., F. Perri, and G.L. Violante (2010) Unequal we stand: An empirical analysis of economic

inequality in the United States, 1967-2006. *Review of Economic Dynamics* 13(1), 15-51.

Heckman, J.J. (1976) A life-cycle model of earnings, learning, and consumption, *Journal of Political Economy* 84, S9-S44.

Jacobs, B. and L. Bovenberg (2011) Optimal taxation of human capital and the earnings function. *Journal of Public Economic Theory* 13(6), 957-971.

James, J. (2012) The college wage premium. *Economic Commentary*. Federal Reserve Bank of Cleveland, Issue 2012-10.

Kapička, M. (2006) Optimal income taxation with human capital accumulation and limited record keeping. *Review of Economic Dynamics* 9, 612–639.

Kapička, M. (2013) Efficient allocations in dynamic private information economies with persistent shocks: a first-order approach. *Review of Economic Studies* 80, 1027–1054.

Kapička, M. (2015) Optimal Mirrleesean taxation in a Ben-Porath economy, *American Economic Journal: Macroeconomics* 7(2), 219-248.

Kapička, M. and J. Neira (2015) Optimal taxation with risky human capital. Working Paper, University of California, Santa Barbara.

Kocherlakota, N.R. (2005) Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation. *Econometrica* 73, 1587-1621.

Kocherlakota, N. R. (2010) *The New Dynamic Public Finance*. Princeton: Princeton University Press.

Kopczuk, W. (2013), Taxation of Intergenerational Transfers and Wealth, in A.J. Auerbach, R. Chetty, M. Feldstein and E. Saez (eds.), *Handbook of Public Economics* 5, 329-390.

Maldonado, D. (2008) Education policies and optimal taxation, *International Tax Public Finance* 15, 131–143.

Manuelli, R.E. and A. Seshadri (2014) Human capital and the wealth of nations. *American Economic Review* 104(9), 2736-2762.

McDaniel, C. (2007) Average tax rates on consumption, investment, labor and capital in the OECD 1950-2003. Working Paper, Arizona State University.

Mirrlees, J. (1971) An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38, 175-208.

Murphy, K.M. and F. Welch (1992) The structure of wages. *Quarterly Journal of Economics* 107(1), 285-326.

Peterman, W.B. (2016) The effect of endogenous human capital accumulation on optimal taxation. *Review of Economic Dynamics* 21, 46-71.

Piketty, T. and E. Saez (2013), Optimal Labor Income Taxation, in A.J. Auerbach, R. Chetty, M. Feldstein and E. Saez (eds.), *Handbook of Public Economics* 5, 391-474.

Saez, E. (2002) Optimal income transfer programs: intensive versus extensive labor supply responses. *Quarterly Journal of Economics* 117, 1039-1073.

Stantcheva, S. (2015) Learning and (or) doing: Human capital investments and optimal taxation. Working Paper, Harvard University.

Stantcheva, S. (2016) Optimal taxation and human capital policies over the life cycle, forthcoming, *Journal of Political Economy*.

Stiglitz, J. (1982) Self-selection and Pareto efficient taxation. *Journal of Public Economics* 17, 213-240.

Todd, P. and K. Wolpin (2003) On the Specification and estimation of the production function for cognitive achievement. *Economic Journal* 113(485), 3–33.

Todd, P. and K. Wolpin (2007) The production of cognitive achievement in children: Home, school and racial test score gaps. *Journal of Human Capital* 1(1), 91-136.

Tuomala, M. (1990), *Optimal Income Taxation and Redistribution.* Oxford, Clarendon Press.

Vogl, T. (2016). Differential fertility, human capital, and development. *Review of Economic Studies* 83, 365–401.

Werning, I. (2007) Optimal fiscal policy with redistribution. *Quarterly Journal of Economics* 122, 925-967.

Appendix A: General Function Form of Technology

In this Appendix, we use a more general function form for the human capital technology $\psi_t^i(h,e)$ and analyze the sign of wedges for the low type in a two-period model.

Assumption 3. Assume that $\psi_t^i(h,e) = \Phi(h,e) + \varepsilon_t^i$, where $\varepsilon_t^H > \varepsilon_t^L$, $\Phi_e(h,e) > 0$ and $\Phi_{ee}(h,e) \le 0$.

Our results depend on neither a specific function form of Φ nor the sign of Φ_{eh}.

Lemma 4. Suppose that Assumptions 1 and 3 hold. Then, the constrained efficient allocation A^* satisfies $e_1^{*L|L} > e_1^{*L|H}$ and $h_2^{*L|L} < h_2^{*L|H}$.

Proof. Denote $d_1^{*L} = c_1^{*L|L} + e_1^{*L|L}$. Then, $c_1^{*L|i} = d_1^{*L} - e_1^{*L|i}$ is derived from $c_1^{*L|H} = c_1^{*L|L} + e_1^{*L|L} - e_1^{*L|H}$.

Based on Assumptions 1 and 3 and Theorem 1, if the reporting type is *L*, then (7a) can be written in a two-period model as follows:

$$u'\left(d_1^{*L} - e_1^{*L|i}\right) - \beta\kappa\gamma\left(z_2^{*L}\right)^\gamma \frac{\Phi_e\left(h_1, e_1^{*L|i}\right)}{\left[\Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i\right]^{\gamma+1}} = 0 \tag{A1}$$

If we differentiate both sides of equation (A1) with respect to ε^i, we get

$$-u''\left(d_1^{*L} - e_1^{*L|i}\right)\frac{de_1^{*L|i}}{d\varepsilon^i} + \beta\kappa\gamma\left(z_2^{*L}\right)^\gamma \left\{ \frac{(\gamma+1)\Phi_e\left(h_1, e_1^{*L|i}\right)}{\left[\Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i\right]^{\gamma+2}} + \frac{de_1^{*L|i}}{d\varepsilon^i}\left[\frac{(\gamma+1)\left[\Phi_e\left(h_1, e_1^{*L|i}\right)\right]^2}{\left[\Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i\right]^{\gamma+2}} - \frac{\Phi_{ee}\left(h_1, e_1^{*L|i}\right)}{\left[\Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i\right]^{\gamma+1}} \right] \right\} = 0,$$

which implies

$$\frac{de_1^{*L|i}}{d\varepsilon^i} = \frac{-\beta\kappa\gamma(\gamma+1)\left(z_2^{*L}\right)^\gamma \frac{\Phi_e\left(h_1, e_1^{*L|i}\right)}{\left[\Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i\right]^{\gamma+2}}}{-u''\left(d_1^{*L} - e_1^{*L|i}\right) + \beta\kappa\gamma\left(z_2^{*L}\right)^\gamma \left[\frac{(\gamma+1)\left[\Phi_e\left(h_1, e_1^{*L|i}\right)\right]^2}{\left[\Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i\right]^{\gamma+2}} - \frac{\Phi_{ee}\left(h_1, e_1^{*L|i}\right)}{\left[\Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i\right]^{\gamma+1}} \right]} \tag{A2}$$

Conditions $u'' < 0$, $\Phi_e > 0$ and $\Phi_{ee} \le 0$ give $\frac{de_1^{*L|i}}{d\varepsilon^i} < 0$. Hence, $\varepsilon^H > \varepsilon^L$ implies $e_1^{*L|L} > e_1^{*L|H}$.

Moreover, $h_2^{*L|i} = \Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i$ leads to $\frac{dh_2^{*L|i}}{d\varepsilon^i} = \Phi_e\left(h_1, e_1^{*L|i}\right)\frac{de_1^{*L|i}}{d\varepsilon^i} + 1$. Using (A2), we get

$$\frac{dh_2^{*L|i}}{d\varepsilon^i} = \frac{-u''\left(d_1^{*L} - e_1^{*L|i}\right) - \beta\kappa\gamma\left(z_2^{*L}\right)^\gamma \frac{\Phi_{ee}\left(h_1, e_1^{*L|i}\right)}{\left[\Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i\right]^{\gamma+1}}}{-u''\left(d_1^{*L} - e_1^{*L|i}\right) + \beta\kappa\gamma\left(z_2^{*L}\right)^\gamma\left[\frac{(\gamma+1)\left[\Phi_e\left(h_1, e_1^{*L|i}\right)\right]^2}{\left[\Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i\right]^{\gamma+2}} - \frac{\Phi_{ee}\left(h_1, e_1^{*L|i}\right)}{\left[\Phi\left(h_1, e_1^{*L|i}\right) + \varepsilon^i\right]^{\gamma+1}}\right]}. \tag{A3}$$

Since $u'' < 0$, $\Phi_{ee} \le 0$, we have $\frac{dh_2^{*L|i}}{d\varepsilon^i} > 0$. Then, $\varepsilon^H > \varepsilon^L$ implies $h_2^{*L|L} < h_2^{*L|H}$. ∎

Proposition 3. Under Assumptions 1 and 3, the intertemporal wedge for the low type is positive, $\tau_{k_2}^L > 0$. The intratemporal wedge for the low type is negative in the first period, $\tau_{z_1}^L < 0$, and positive in the second period, $\tau_{z_2}^L > 0$.

Proof. Since the human capital investment is zero for everyone in the terminal period, the fact that $c_t^{*L|H} = c_t^{*L|L} + e_t^{*L|L} - e_t^{*L|H}$, $t = 1, 2$, and Lemma 4 imply that

$$1 = \frac{u'\left(c_2^{*L|H}\right)}{u'\left(c_2^{*L|L}\right)} > \frac{u'\left(c_1^{*L|H}\right)}{u'\left(c_1^{*L|L}\right)}, \tag{A4}$$

and

$$1 = \frac{u'\left(c_2^{*L|H}\right)}{u'\left(c_2^{*L|L}\right)} > \frac{v'\left(\frac{z_2^{*L}}{h_2^{*L|H}}\right)\frac{1}{h_2^{*L|H}}}{v'\left(\frac{z_2^{*L}}{h_2^{*L|L}}\right)\frac{1}{h_2^{*L|L}}}. \tag{A5}$$

By (6a), (A4) implies a positive intertemporal wedge for the low type, $\tau_{k_2}^L > 0$. Moreover, by (6b), (A5) implies that the intratemporal wedge for the low type is positive in the second period, $\tau_{z_2}^L > 0$. By contrast, in the first period, $v'\left(\frac{z_1^{*L}}{h_1^{*L|H}}\right)\frac{1}{h_1^{*L|H}} = v'\left(\frac{z_1^{*L}}{h_1^{*L|L}}\right)\frac{1}{h_1^{*L|L}}$, since, by construction, initial human capital is the same for the both types. Then, $e_1^{*L|L} > e_1^{*L|H}$ in Lemma 4 implies that

$$\frac{u'\left(c_1^{*L|H}\right)}{u'\left(c_1^{*L|L}\right)} < \frac{v'\left(\frac{z_1^{*L}}{h_1^{*L|H}}\right)\frac{1}{h_1^{*L|H}}}{v'\left(\frac{z_1^{*L}}{h_1^{*L|L}}\right)\frac{1}{h_1^{*L|L}}} = 1. \tag{A6}$$

By (6b), (A6) implies that the intratemporal wedge for the low type is negative in the first period, $\tau_{z_1}^L < 0$. ∎

Appendix B

This Appendix offers proofs for the lemmas, propositions and theorems stated in the text.

Proof of Theorem 1.

Since it takes one period for the human capital investment to accumulate human capital, the human capital investment in the terminal period T is useless. Hence, the optimal allocation of $e_T^{*r|i}$ is zero.

The Lagrangian of the problem described in Section 3.3 is set as follows:

$$\mathscr{L} = \max_{\{c_t,e_t\}_{t=1}^{T},\{h_t\}_{t=2}^{T}} \sum_{t=1}^{T} \beta^{t-1}\left[u(c_t)-v\left(\tfrac{z_t^{*r}}{h_t}\right)\right]+\lambda_t\left\{c_t^{*r|r}+e_t^{*r|r}-c_t-e_t\right\}+\mu_t\left\{\psi_t^i(h_t,e_t)-h_{t+1}\right\}.$$

The first-order conditions with respect to $\{c_t,e_t\}_{t=1}^{T}$, $\{h_t\}_{t=2}^{T}$ are as follows:

$$\beta^{t-1}u'(c_t)=\lambda_t, \quad t=1,2,...T; \tag{B1a}$$

$$\lambda_t = \mu_t \tfrac{\partial}{\partial e}\left[\psi_t^i(h_t,e_t)\right], \quad t=1,2,...T; \tag{B1b}$$

$$\mu_{t-1} = \mu_t \tfrac{\partial}{\partial h}\left[\psi_t^i(h_t,e_t)\right]+\beta^{t-1}v'\left(\tfrac{z_t^{*r}}{h_t}\right)\tfrac{z_t^{*r}}{(h_t)^2}, \quad t=2,3,...T-1; \tag{B1c}$$

$$\mu_{T-1} = \beta^{T-1}v'\left(\tfrac{z_T^{*r}}{h_T}\right)\tfrac{z_T^{*r}}{(h_T)^2}. \tag{B1d}$$

While (B1a) and (B1b) imply

$$\mu_t = \frac{\beta^{t-1}u'(c_t)}{\tfrac{\partial}{\partial e}\left[\psi_t^i(h_t,e_t)\right]}, \quad t=1,2,...T; \tag{B2a}$$

(B1c) and (B2a) lead to

$$\frac{\beta^{t-2}u'(c_{t-1})}{\tfrac{\partial}{\partial e}\left[\psi_{t-1}^i(h_{t-1},e_{t-1})\right]} = \frac{\beta^{t-1}u'(c_t)}{\tfrac{\partial}{\partial e}\left[\psi_t^i(h_t,e_t)\right]}\tfrac{\partial}{\partial h}\left[\psi_t^i(h_t,e_t)\right]+\beta^{t-1}v'\left(\tfrac{z_t^{*r}}{h_t}\right)\tfrac{z_t^{*r}}{(h_t)^2}, \quad t=2,3,...T-1; \tag{B2b}$$

and (B1a), (B1b) and (B1d) result in

$$u'(c_{T-1}) = \beta v'\left(\tfrac{z_T^{*r}}{h_T}\right)\tfrac{z_T^{*r}}{(h_T)^2}\tfrac{\partial}{\partial e}\left[\psi_{T-1}^i(h_{T-1},e_{T-1})\right], \tag{B2c}$$

which is (7a) in the text.

Moreover, using (B2b), (B2c) at $t=T$-1 gives

$$\frac{u'(c_{T-2})}{\tfrac{\partial}{\partial e}\left[\psi_{T-2}^i(h_{T-2},e_{T-2})\right]} = \beta^2 v'\left(\tfrac{z_T^{*r}}{h_T}\right)\tfrac{z_T^{*r}}{(h_T)^2}\tfrac{\partial}{\partial h}\left[\psi_{T-1}^i(h_{T-1},e_{T-1})\right]+\beta v'\left(\tfrac{z_{T-1}^{*r}}{h_{T-1}}\right)\tfrac{z_{T-1}^{*r}}{(h_{T-1})^2}. \tag{B2d}$$

We repeat the process for periods $t=T$-2 to $t=1$ by using (B2b) and (B2d). For example, we rewrite (B2b) for $t=T$-2, and substitute the term $u'(c_{T-2})$ in the left-hand side of the resulting equation by $u'(c_{T-2})$ in (B2d). We get an expression for $t=T$-3 like (B2d). We repeat the same process for $t=T$-3, …,1. Then, we obtain

$$\frac{u'(c_t)}{\tfrac{\partial}{\partial e}\left[\psi_t^i(h_t,e_t)\right]} = \frac{\beta u'(c_{t+1})}{\tfrac{\partial}{\partial e}\left[\psi_{t+1}^i(h_{t+1},e_{t+1})\right]}\tfrac{\partial}{\partial h}\left[\psi_{t+1}^i(h_{t+1},e_{t+1})\right]+\beta v'\left(\tfrac{z_{t+1}^{*r}}{h_{t+1}}\right)\tfrac{z_{t+1}^{*r}}{(h_{t+1})^2}$$

$$= \beta\left\{v'\left(\tfrac{z_{t+1}^{*r}}{h_{t+1}}\right)\tfrac{z_{t+1}^{*r}}{(h_{t+1})^2}+\sum_{q=t+2}^{T}\left[\beta^{q-t-1}v'\left(\tfrac{z_q^{*r}}{h_q}\right)\tfrac{z_q^{*r}}{(h_q)^2}\prod_{s=t+1}^{q-1}\tfrac{\partial}{\partial h}\left[\psi_s^i(h_s,e_s)\right]\right]\right\}, \quad t=1,...,T-1, \tag{B2e}$$

which is (7b) in the text.

Therefore, the constrained efficient allocation $\{c_t^{*r|i},e_t^{*r|i},h_t^{*r|i}\}_{t=1}^{T}$ must satisfy (7a) and (7b). ∎

Next, we proceed to prove Lemmas 1, 2 and 3. Based on Assumptions 1 and 2 and given that agents report as low type $r=L$, (7a) and (7b) can be written as follows

$$u'\left(c_{T-1}^{*L|i}\right) = \alpha\beta\gamma\kappa\left(z_T^{*L}\right)^{\gamma}\left(h_T^{*L|i}\right)^{-\gamma-1}, \tag{B3a}$$

$$u'\left(c_t^{*L|i}\right)=\alpha\beta\kappa\gamma\sum_{q=t+1}^{T}\left\{\left[(1-\delta_h)\beta\right]^{q-t-1}\left(z_q^{*L}\right)^{\gamma}\left(h_q^{*L|i}\right)^{-\gamma-1}\right\},\quad t=1,2,...,T-1. \tag{B3b}$$

These two equations together imply

$$u'\left(c_{t-1}^{*L|i}\right)=\alpha\beta\kappa\gamma\left(z_t^{*L}\right)^{\gamma}\left(h_t^{*L|i}\right)^{-\gamma-1}+\beta(1-\delta_h)u'\left(c_t^{*L|i}\right),\ t=2,3...,T. \tag{B3c}$$

Using the above three equations, we can prove Lemmas 1-3 as follows.

Proof of Lemma 1.

We prove the lemma by contradiction. Suppose that the statement in Lemma 1 is not true. That is,

$$u'\left(c_{T-1}^{*L|L}\right)\le u'\left(c_{T-1}^{*L|H}\right). \tag{B4a}$$

Based on the facts that $c_{T-1}^{*L|H}=c_{T-1}^{*L|L}+e_{T-1}^{*L|L}-e_{T-1}^{*L|H}$ and u is strictly concave, inequality (B4a) implies that $e_{T-1}^{*L|L}\le e_{T-1}^{*L|H}$. Moreover, using (B3a), inequality (B4a) implies $h_T^{*L|H}\le h_T^{*L|L}$. By the function form $\psi_{T-1}^i(h_{T-1},e_{T-1})=(1-\delta_h)h_{T-1}+\alpha e_{T-1}$, conditions $e_{T-1}^{*L|L}\le e_{T-1}^{*L|H}$ and $h_T^{*L|H}\le h_T^{*L|L}$ imply $h_{T-1}^{*L|H}\le h_{T-1}^{*L|L}$. Using (B3b), conditions $h_{T-1}^{*L|H}\le h_{T-1}^{*L|L}$ and $h_T^{*L|H}\le h_T^{*L|L}$ imply that the inequality (B4a) also holds in $t=T$-2:

$$u'\left(c_{T-2}^{*L|L}\right)\le u'\left(c_{T-2}^{*L|H}\right). \tag{B4b}$$

Repeating the above procedure from time $t=T$-2 to $t=2$, we finally get

$$u'\left(c_2^{*L|L}\right)\le u'\left(c_2^{*L|H}\right),\ e_1^{*L|L}\le e_1^{*L|H}\quad\text{and}\quad h_2^{*L|H}\le h_2^{*L|L}.$$

By the functional form $h_2^{*L|i}=\psi_1^i\left(h_1^{*L|i},e_1^{*L|i}\right)=(1-\delta_h)h_1^{*L|i}+\alpha e_1^{*L|i}+\varepsilon_1^i$, conditions $e_1^{*L|L}\le e_1^{*L|H}$ and $h_2^{*L|H}\le h_2^{*L|L}$ and the initial condition $h_1^{*L|H}=h_1^{*L|L}$ together imply $\varepsilon_1^H\le\varepsilon_1^L$, which is contradictory to the fact that $\varepsilon_1^H>\varepsilon_1^L$. Hence, we complete the proof. ∎

Proof of Lemma 2.

By Lemma 1, (B3a) implies $h_T^{*L|H}>h_T^{*L|L}$. Also, the facts that u is strictly concave and $c_{T-1}^{*L|H}=c_{T-1}^{*L|L}+e_{T-1}^{*L|L}-e_{T-1}^{*L|H}$ imply

$$e_{T-1}^{*L|L}>e_{T-1}^{*L|H}. \tag{B5a}$$

By the functional form $\psi_{T-1}^i\left(h_{T-1}^{*L|i},e_{T-1}^{*L|i}\right)=(1-\delta_h)h_{T-1}^{*L|i}+\alpha e_{T-1}^{*L|i}$, conditions $e_{T-1}^{*L|L}>e_{T-1}^{*L|H}$ and $h_T^{*L|H}>h_T^{*L|L}$ imply $h_{T-1}^{*L|H}>h_{T-1}^{*L|L}$. Using condition $h_q^{*L|H}>h_q^{*L|L}$ for $q\ge T-1$, (B3b) implies $u'\left(c_{T-2}^{*L|L}\right)>u'\left(c_{T-2}^{*L|H}\right)$. Condition $u'\left(c_{T-2}^{*L|L}\right)>u'\left(c_{T-2}^{*L|H}\right)$ and the facts that u is strictly concave and $c_{T-2}^{*L|H}=c_{T-2}^{*L|L}+e_{T-2}^{*L|L}-e_{T-2}^{*L|H}$ together imply that the inequality (B5a) also holds in $t=T$-2:

$$e_{T-2}^{*L|L}>e_{T-2}^{*L|H}. \tag{B5b}$$

Repeating the above procedure from time $t=T$-2 to $t=2$, we obtain

$$e_t^{*L|L}>e_t^{*L|H}\quad\text{for}\quad t=1,...,T-1\quad\text{and}\quad h_t^{*L|L}<h_t^{*L|H}\quad\text{for}\quad t=2,...,T.$$

Note that the relationships $h_t^{*L|L}<h_t^{*L|H}$ and $e_t^{*L|L}>e_t^{*L|H}$ for $t=2,...,T-1$ imply

32

$$\frac{e_t^{*L|L}}{h_t^{*L|L}} > \frac{e_t^{*L|H}}{h_t^{*L|H}}, \quad t = 2,...,T-1. \tag{B6a}$$

Since $h_{t+1} = \psi_t^i(h_t, e_t) = (1-\delta)h_t + \alpha e_t$ for $t = 2,...,T-1$, we get

$$\frac{h_{t+1}^{*L|L}}{h_{t+1}^{*L|H}} = \frac{(1-\delta_h)h_t^{*L|L} + \alpha e_t^{*L|L}}{(1-\delta_h)h_t^{*L|H} + \alpha e_t^{*L|H}} = \frac{1-\delta_h + \alpha \frac{e_t^{*L|L}}{h_t^{*L|L}}}{1-\delta_h + \alpha \frac{e_t^{*L|H}}{h_t^{*L|H}}} \frac{h_t^{*L|L}}{h_t^{*L|H}}. \tag{B6b}$$

By the inequality (B6a), we have $1-\delta_h + \alpha \frac{e_t^{*L|L}}{h_t^{*L|L}} > 1-\delta_h + \alpha \frac{e_t^{*L|H}}{h_t^{*L|H}}$. Therefore, (B6b) implies

$$\frac{h_t^{*L|L}}{h_t^{*L|H}} < \frac{h_{t+1}^{*L|L}}{h_{t+1}^{*L|H}}$$

for any $t = 2,...,T-1$. ∎

Proof of Lemma 3.

By Theorem 1, $e_T^{*L|L} = e_T^{*L|H} = 0$. Then, the constraint $c_T^{*L|H} = c_T^{*L|L} + e_T^{*L|L} - e_T^{*L|H}$ implies $\frac{c_T^{*L|H}}{c_T^{*L|L}} = 1$.

We prove Lemma 3 by contradiction. Suppose the statement is not true. That is, there exists a $t \in \{2,3,...T\}$ such that

$$\frac{u'\left(c_{t-1}^{*L|H}\right)}{u'\left(c_{t-1}^{*L|L}\right)} \geq \frac{u'\left(c_t^{*L|H}\right)}{u'\left(c_t^{*L|L}\right)}. \tag{B7a}$$

Using (B3c), inequality (B7a) is written as

$$\frac{\alpha\beta\kappa\gamma\left(z_t^{*L}\right)^\gamma\left(h_t^{*L|H}\right)^{-\gamma-1} + \beta(1-\delta_h)u'\left(c_t^{*L|H}\right)}{\alpha\beta\kappa\gamma\left(z_t^{*L}\right)^\gamma\left(h_t^{*L|L}\right)^{-\gamma-1} + \beta(1-\delta_h)u'\left(c_t^{*L|L}\right)} \geq \frac{u'\left(c_t^{*L|H}\right)}{u'\left(c_t^{*L|L}\right)}, \tag{B7b}$$

which implies

$$\left(h_t^{*L|H}\right)^{-\gamma-1}u'\left(c_t^{*L|L}\right) \geq \left[h_t^{*L|L}\right]^{-\gamma-1}u'\left(c_t^{*L|H}\right). \tag{B7c}$$

Moreover, using (B3b), inequality (B7c) is written as

$$\alpha\beta\kappa\gamma\sum_{q=t+1}^{T}\left\{\left[(1-\delta_h)\beta\right]^{q-t-1}\left(z_q^{*L}\right)^\gamma\left(\frac{h_t^{*L|L}}{h_q^{*L|L}}\right)^{\gamma+1}\right\} \geq \alpha\beta\kappa\gamma\sum_{q=t+1}^{T}\left\{\left[(1-\delta_h)\beta\right]^{q-t-1}\left(z_q^{*L}\right)^\gamma\left(\frac{h_t^{*L|H}}{h_q^{*L|H}}\right)^{\gamma+1}\right\}. \tag{B7d}$$

By contrast, by Lemma 2, we have $\frac{h_t^{*L|L}}{h_t^{*L|H}} < \frac{h_q^{*L|L}}{h_q^{*L|H}}$ for $2 \leq t < q$, which implies

$$\frac{h_t^{*L|L}}{h_q^{*L|L}} < \frac{h_t^{*L|H}}{h_q^{*L|H}} \quad \text{for } 2 \leq t < q. \tag{B8a}$$

Hence, based on inequality (B8a), we can derive

$$\alpha\beta\kappa\gamma\sum_{q=t+1}^{T}\left\{\left[(1-\delta_h)\beta\right]^{q-t-1}\left(z_q^{*L}\right)^\gamma\left(\frac{h_t^{*L|L}}{h_q^{*L|L}}\right)^{\gamma+1}\right\} < \alpha\beta\kappa\gamma\sum_{q=t+1}^{T}\left\{\left[(1-\delta_h)\beta\right]^{q-t-1}\left(z_q^{*L}\right)^\gamma\left(\frac{h_t^{*L|H}}{h_q^{*L|H}}\right)^{\gamma+1}\right\} \tag{B8b}$$

which is contradictory to inequality (B7d). ∎

Proof of Proposition 2.

As all agents have the same human capital $h_1^{*L|L} = h_1^{*L|H} = h_1^{*H|H} = h_1$ in the first period, we have

$v'\left(\frac{z_1^{*L}}{h_1^{*L|H}}\right)\frac{1}{h_1^{*L|H}}=v'\left(\frac{z_1^{*L}}{h_1^{*L|L}}\right)\frac{1}{h_1^{*L|L}}$. By Lemma 2, $e_1^{*L|H}<e_1^{*L|L}$, which implies $\frac{u'\left(c_1^{*L|H}\right)}{u'\left(c_1^{*L|L}\right)}=\frac{u'\left(c_1^{*L|L}+e_1^{*L|L}-e_1^{*L|H}\right)}{u'\left(c_1^{*L|L}\right)}<1.$

Therefore, (6b) and the fact $\frac{u'\left(c_1^{*L|H}\right)}{u'\left(c_1^{*L|L}\right)}<\frac{v'\left(z_1^{*L}/h_1^{*L|H}\right)\frac{1}{h_1^{*L|H}}}{v'\left(z_1^{*L}/h_1^{*L|L}\right)\frac{1}{h_1^{*L|L}}}=1$ imply that $\tau_{z_1}^L<0.$

Moreover, according to Theorem 1, $e_T^{*L|L}=e_T^{*L|H}=0$, thus implying $u'\left(c_T^{*L|H}\right)=u'\left(c_T^{*L|L}\right)$. According to Lemma 2, $h_T^{*L|H}>h_T^{*L|L}$. Assumption 1 suggests $v'\left(\frac{z_T^{*L}}{h_T^{*L|H}}\right)\frac{1}{h_T^{*L|H}}<v'\left(\frac{z_T^{*L}}{h_T^{*L|L}}\right)\frac{1}{h_T^{*L|L}}$. As a

consequence of (6b) and the fact $\frac{v'\left(z_T^{*L}/h_T^{*L|H}\right)\frac{1}{h_T^{*L|H}}}{v'\left(z_T^{*L}/h_T^{*L|L}\right)\frac{1}{h_T^{*L|L}}}<\frac{u'\left(c_T^{*L|H}\right)}{u'\left(c_T^{*L|L}\right)}=1$, we obtain $\tau_{z_T}^L>0.$ ∎

Proof of Theorem 2.

To prove that the tax system $\{T_t^{e*}(k^t,z^t)\}_{t=1}^T$ is optimal, we need to show that the constrained efficient allocation A^* is a competitive equilibrium. That is, A^* must satisfy the following three conditions:
(a) consistent with competitive pricing conditions at prices $\{r_t^*,w_t^*\}_{t=1}^T$,
(b) consistent with the market clearing condition;
(c) consistent with agent's utility maximization, under tax system $\{T_t^{e*}(k^t,z^t)\}_{t=1}^T$ and prices $\{r_t^*,w_t^*\}_{t=1}^T$.

For (a), from the conditions of the firm's problem in subsection 5.2, we get

$$w_t^*=F_z\left(K_t^*,Z_t^*\right),$$

$$r_t^*=F_k\left(K_t^*,Z_t^*\right)-\delta_k.$$

For (b), as the constrained efficient allocation A^* is feasible, it satisfies the market clearing condition.

What remains to be shown is that, given the tax system $\{T_t^{e*}(k^t,z^t)\}_{t=1}^T$ and prices $\{r_t^*,w_t^*\}_{t=1}^T$, the tuple $(c^*,e^*,h^*,l^*,z^*,k^*)$ is individually optimal for each agent.

Given the tax system stated in subsection 5.5 with the following lump-sum taxes and parameter ρ_t^i in the condition $\rho_t^i\left(k_t-k_t^{*i}\right)=\left(z_t^{*i}-z_t\right)$, $i=L,H$, we now show that the constrained efficient allocation A^* is indeed implemented by this tax system:

$$\Gamma_t^L=\frac{v'\left(\frac{z_t^{*L}}{h_t^{*L|L}}\right)\frac{z_t^{*L}}{h_t^{*L|L}}+u'\left(c_{t-1}^{*L|L}\right)\frac{k_t^{*L}}{\beta}}{u'\left(c_t^{*L|L}\right)}-c_t^{*L|L}-e_t^{*L|L}-k_{t+1}^{*L},\tag{B9a}$$

$$\Gamma_t^H=\frac{v'\left(\frac{z_t^{*H}}{h_t^{*H|H}}\right)\frac{z_t^{*H}}{h_t^{*H|H}}+u'\left(c_{t-1}^{*H|H}\right)\frac{k_t^{*H}}{\beta}}{u'\left(c_t^{*H|H}\right)}-c_t^{*H|H}-e_t^{*H|H}-k_{t+1}^{*H}.\tag{B9b}$$

34

$$\rho_t^L = \frac{u'\left(c_t^{*L|H}\right)u'\left(c_{t-1}^{*L|L}\right) - u'\left(c_t^{*L|L}\right)u'\left(c_{t-1}^{*L|H}\right)}{\beta\left[u'\left(c_t^{*L|H}\right)\frac{1}{h_t^{*L|L}}v'\left(\frac{z_t^{*L}}{h_t^{*L|L}}\right) - u'\left(c_t^{*L|L}\right)\frac{1}{h_t^{*L|H}}v'\left(\frac{z_t^{*L}}{h_t^{*L|H}}\right)\right]}, \qquad t=1,\,2,\,T\,.$$ (B9c)

$$\rho_t^H = 0, \qquad t=1,\,2,\,T\,.$$ (B9d)

Due to the fact that the tax system $\{T_t^{e*}(k^t, z^t)\}_{t=1}^T$ severely punishes any observable allocation (z^T, k^T) which satisfies either $\rho_t^L\left(k_t - k_t^{*L}\right) = \left(z_t^{*L} - z_t\right)$ or $\rho_t^H\left(k_t - k_t^{*H}\right) = \left(z_t^{*H} - z_t\right)$ for $t \leq T$, an agent will only consider the following two strategies:

Strategy L: choose (z^T, k^T) such that $\rho_t^L\left(k_t - k_t^{*L}\right) = \left(z_t^{*L} - z_t\right)$ for any $t \leq T$; or

Strategy H: choose (z^T, k^T) such that $\rho_t^H\left(k_t - k_t^{*H}\right) = \left(z_t^{*H} - z_t\right)$ for any $t \leq T$.

Suppose that a type i agent chooses Strategy r, $r \in \{L, H\}$. She will choose the optimal allocation (c, e, h, l, k) to solve the following problem:

$$\max_{c_t, e_t, h_t, l_t, k_{t+1}} \sum_{t=1}^T \beta^{t-1}\left[u(c_t) - v(l_t)\right]$$

subject to the budget constraints

$$c_t + e_t + k_{t+1} + \Gamma_t^r \leq \left(1 - \tilde{\tau}_{z_t}^r\right)w_t^* h_t l_t + \left(1 - \tilde{\tau}_{k_t}^r\right)\left(1 + r_t^*\right)k_t \text{ for } t = 1, 2..., T,$$ (B10a)

the strategy restriction

$$\rho_t^r\left(k_t - k_t^{*r}\right) + \left(l_t h_t - z_t^{*r}\right) = 0 \text{ for } t = 1, 2, ..., T,$$ (B10b)

and the law of motion of human capital

$$h_{t+1} = \psi_t^i(h_t, e_t) \quad \text{for } t = 1, 2, ..., T-1,$$ (B10c)

where h_1 and k_1 are given. We refer to the above decentralized problem as $P^{r|i}$.

Let λ_t, η_t and μ_t be the multiplier for (B10a), (B10b) and (B10c), respectively. The Lagrangian is

$$\begin{aligned}
\mathscr{L} = \max_{c_t, l_t, k_{t+1}, e_t, h_t} &\sum_{t=1}^T \beta^{t-1}\left[u(c_t) - v(l_t)\right] \\
&+ \lambda_t\left[\left(1 - \tilde{\tau}_{z_t}^r\right)w_t^* h_t l_t + \left(1 - \tilde{\tau}_{k_t}^r\right)\left(1 + r_t^*\right)k_t - c_t - e_t - k_{t+1} - \Gamma_t^r\right] \\
&+ \eta_t\left[\rho_t^r\left(k_t^{r*} - k_t\right) + \left(z_t^{r*} - l_t h_t\right)\right] \\
&+ \mu_t\left[\psi_t^i(h_t, e_t) - h_{t+1}\right]
\end{aligned}$$

The first-order conditions are

$[c_t]$: $\qquad \beta^{t-1}u'(c_t) - \lambda_t = 0, \quad t = 1, ..., T,$

$[l_t]$: $\qquad -\beta^{t-1}v'(l_t) + \lambda_t w_t^* h_t\left(1 - \tilde{\tau}_{z_t}^r\right) - \eta_t h_t = 0, \quad t = 1, ..., T,$

$[k_t]$: $\qquad \lambda_t\left(1 + r_t^*\right)\left(1 - \tilde{\tau}_{k_t}^r\right) - \lambda_{t-1} - \eta_t \rho_t^r = 0, \quad t = 2, ..., T,$

$[e_t]$: $\qquad -\lambda_t + \mu_t \frac{\partial}{\partial e}\left[\psi_t^i(h_t, e_t)\right] = 0, \quad t = 1, ..., T-1,$

$[h_t]$: $\qquad \lambda_t\left(1 - \tilde{\tau}_{z_t}^r\right)w_t^* l_t - \eta_t l_t + \mu_t \frac{\partial}{\partial h}\left[\psi_t^i(h_t, e_t)\right] - \mu_{t-1} = 0, \quad t = 2, ..., T-1,$

$[h_T]$: $\qquad \lambda_T\left(1 - \tilde{\tau}_{\tilde{z}}\right)w^* l \mp \eta l_T \mp \mu_{1\ T} = 0$

Using the same method as the Proof of Theorem 1, these first-order conditions imply the following three equations

$$u'(c_{T-1}) = \beta v'(l_T)\frac{l_T}{h_T}\frac{\partial}{\partial h}\left[\psi_{T-1}^i(h_{T-1}, e_{T-1})\right], \tag{B11a}$$

$$\frac{u'(c_t)}{\frac{\partial}{\partial e}\left[\psi_t^i(h_t, e_t)\right]} = \left\{\beta v'(l_{t+1})\frac{l_{t+1}}{h_{t+1}} + \sum_{q=t+1}^{T-1}\beta^{q-t+1}v'(l_{q+1})\frac{l_{q+1}}{h_{q+1}}\prod_{s=t+1}^{q}\frac{\partial}{\partial h}\left[\psi_s^i(h_s, e_s)\right]\right\}, \tag{B11b}$$

$$u'(c_{t-1}) + \beta\left[u'(c_t)w_t^*\left(1 - \tilde{\tau}_{z_t}^r\right) - \frac{1}{h_t}v'(l_t)\right]\rho_t^r = \beta u'(c_t)\left(1 + r_t^*\right)\left(1 - \tilde{\tau}_{k_t}^r\right). \tag{B11c}$$

Although the first-order conditions are necessary and not sufficient, it is noted that our problem uses a quadratic objective function with linear constraints. Then, these first-order conditions are also sufficient if the Jacobean matrix of the first-order conditions are negative semi-definite. Denote $\nabla^2 \mathscr{L}$ as the Jacobean matrix of the first-order conditions. See the Jacobean matrix in Appendix Table. Obviously, the matrix is negative semi-definite, since one of the columns has all elements equal to zero. Hence, there is a unique solution to the above problem. With the Inada conditions, the solution is interior. Therefore, the first-order conditions are both necessary and sufficient for the maximum. Now, we show that for agents who choose strategy $\sigma = (r|i)$, if their constrained efficient allocation satisfies all (B10a)-(B10c) and (B11a)-(B11c), then the allocation must be the unique solution of the above decentralized problem.

For any strategy $(r|i) \in \{(H|H),(L|L),(L|H)\}$, by (10a)-(10c) and (B9a)-(B9b), we know that the tuple $\{c_t^{*r|i}, e_t^{*r|i}, h_t^{*r|i}, \frac{z_t^{*r}}{h_t^{*r|i}}, k_t^{*r}\}_{t=1}^T$ satisfies (B10a) written as an equality. Also, the tuple $\{e_t^{*r|i}, h_t^{*r|i}, \frac{z_t^{*r}}{h_t^{*r|i}}, k_t^{*r}\}_{t=1}^T$ satisfies (B10b) and (B10c). By Theorem 1, the tuple $\{c_t^{*r|i}, e_t^{*r|i}, h_t^{*r|i}, \frac{z_t^{*r}}{h_t^{*r|i}}, k_t^{*r}\}_{t=1}^T$ satisfies (B11a) and (B11b). By (10a) and (10b), the tuple $\{c_t^{*r|i}, e_t^{*r|i}, h_t^{*r|i}, \frac{z_t^{*r}}{h_t^{*r|i}}, k_t^{*r}\}_{t=1}^T$ satisfies (B11c). Therefore the constrained efficient allocation $\{c_t^{*r|i}, e_t^{*r|i}, h_t^{*r|i}, \frac{z_t^{*r}}{h_t^{*r|i}}, k_t^{*r}\}_{t=1}^T$ is indeed the unique solution of Problem $P^{r|i}$.

Consider a high-type agent. If she chooses Strategy H, then she solves the problem $P^{H|H}$ and the resulting individual optimal allocation is $\{c_t^{*H|H}, e_t^{*H|H}, h_t^{*H|H}, \frac{z_t^{*H}}{h_t^{*H|H}}, k_t^{*H}\}_{t=1}^T$. By contrast, if she chooses Strategy L, she solves the problem $P^{L|H}$ and the resulting individual optimal allocation is $\{c_t^{*L|H}, e_t^{*L|H}, h_t^{*L|H}, \frac{z_t^{*L}}{h_t^{*L|H}}, k_t^{*L}\}_{t=1}^T$. The incentive compatibility constraint implies that the allocation does not

give a higher utility than the allocation when Strategy H is chosen. Therefore, the best strategy for high types is Strategy H and the constrained efficient allocation $\{c_t^{*H|H}, e_t^{*H|H}, h_t^{*H|H}, \frac{z_t^{*H}}{h_t^{*H|H}}, k_t^{*H}\}_{t=1}^T$ indeed solves the individual utility maximization problem for high types.

Next, consider a low-type agent. If she chooses Strategy L, then she solves the problem $P^{L|L}$ and the individual optimal allocation is $\{c_t^{*L|L}, e_t^{*L|L}, h_t^{*L|L}, \frac{z_t^{*L}}{h_t^{*L|L}}, k_t^{*L}\}_{t=1}^T$. By contrast, if she chooses Strategy H, then she must provide effective labor z_t^{*H} for any $t \leq T$. When the difference between these two types is large enough, e.g., $\bar{l} < \frac{z_2^{*H}}{\psi_1^L(h_1, \bar{e}_1)}$, the low-type agent cannot choose Strategy H. Therefore, the best strategy for a low type is Strategy L and thus, the constrained efficient allocation $\{c_t^{*L|L}, e_t^{*L|L}, h_t^{*L|L}, \frac{z_t^{*L}}{h_t^{*L|L}}, k_t^{*L}\}_{t=1}^T$ indeed solves the individual utility maximization problem for a low type. Hence, the constrained efficient allocation A^* is indeed a competitive equilibrium. ∎

Appendix C: First-Order Approach

In this Appendix, we prove that the results obtained in the text with discrete-type agents are robust under the first-order approach. Now, there are continuous-type agents. Let an agent's ability be indexed by ε that lies within the interval $[\underline{\varepsilon}, \overline{\varepsilon}]$. We study the first-order approach in a simple model with two periods. We show that, except for the agents of the top type $\overline{\varepsilon}$ and the bottom type $\underline{\varepsilon}$, the capital wedge is positive and the labor wedge is negative in the first period and positive in the second period. The environment is as follows.

Technology: All agents have the same initial human capital h_1, but they have different technology $\psi^\varepsilon(h_1, e_1)$ which satisfies:

Assumption 4. $\psi^\varepsilon(h_1, e_1) = \Phi(h_1, e_1) + \varepsilon$, where $\Phi_e > 0$ and $\Phi_{ee}(h, e) \leq 0$.

Social welfare function: The lifetime utility of an agent with type ε who reports as type $\hat{\varepsilon}$ is:

$$W(\hat{\varepsilon} \mid \varepsilon) = u\left(d_1(\hat{\varepsilon}) - e_1(\hat{\varepsilon} \mid \varepsilon)\right) - v\left(\frac{z_1(\hat{\varepsilon})}{h_1}\right) + \beta\left[u\left(d_2(\hat{\varepsilon})\right) - v\left(\frac{z_2(\hat{\varepsilon})}{\Phi(h_1, e_1(\hat{\varepsilon} \mid \varepsilon)) + \varepsilon}\right)\right].$$

To save the notation and without loss of generality, $W(\varepsilon|\varepsilon)$ will be written as $W(\varepsilon)$. Let $q(\varepsilon)$ be the probability distribution of types. Then, the social welfare is written as follows:

$$\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} W(\varepsilon) q(\varepsilon) d\varepsilon.$$

Incentive compatible constraint: By the revelation principle, we restrict attention to the allocation that satisfies the following incentive compatible constraint:

$$W(\varepsilon \mid \varepsilon) \geq W(\hat{\varepsilon} \mid \varepsilon) \quad \text{for any} \quad \hat{\varepsilon} \in [\underline{\varepsilon}, \overline{\varepsilon}]$$

Therefore, the incentive compatible constraint can be written as

$$W(\varepsilon) = \max_{\hat{\varepsilon}} W(\hat{\varepsilon} \mid \varepsilon).$$

The envelope condition is

$$\dot{W}(\varepsilon) = -u'\big(d_1(\varepsilon) - e_1(\varepsilon \mid \varepsilon)\big) D_2 e_1(\varepsilon \mid \varepsilon) + \beta v'\left(\frac{z_2(\varepsilon)}{\Phi(h_1, e_1(\varepsilon \mid \varepsilon)) + \varepsilon}\right) \frac{z_2(\varepsilon)\big[\Phi_e(h_1, e_1(\varepsilon \mid \varepsilon)) D_2 e_1(\varepsilon \mid \varepsilon) + 1\big]}{\big[\Phi(h_1, e_1(\varepsilon \mid \varepsilon)) + \varepsilon\big]^2},$$

where $D_2 e_1(\varepsilon \mid \varepsilon)$ is the derivative of $e_1(\varepsilon \mid \varepsilon)$ with respect to its second argument.

In its integral version, the envelope condition is rewritten as

$$W(\varepsilon) = W(\underline{\varepsilon}) - \int_{\underline{\varepsilon}}^{\varepsilon} \left\{ \begin{aligned} & u'\big(d_1(\hat{\varepsilon}) - e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon})\big) D_2 e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon}) \\ & - \beta v'\left(\frac{z_2(\tilde{\varepsilon})}{\Phi(h_1, e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon})) + \tilde{\varepsilon}}\right) \frac{z_2(\tilde{\varepsilon})\big[\Phi_e(h_1, e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon})) D_2 e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon}) + 1\big]}{\big[\Phi(h_1, e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon})) + \tilde{\varepsilon}\big]^2} \end{aligned} \right\} q(\tilde{\varepsilon}) d\tilde{\varepsilon}.$$

Resource constraints: A feasible allocation must satisfy the resource constraints in each period:

$$\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} d_1(\varepsilon) q(\varepsilon) d\varepsilon + K_2 \leq F\left(K_1, \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} z_1(\varepsilon) q(\varepsilon) d\varepsilon\right) + (1 - \delta_k) K_1 - G_1,$$

$$\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} d_2(\varepsilon) q(\varepsilon) d\varepsilon \leq F\left(K_2, \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} z_2(\varepsilon) q(\varepsilon) d\varepsilon\right) + (1 - \delta_k) K_2 - G_2.$$

Lagrangian: Let λ_1 and λ_2 be the Lagrange multipliers of the two resource constraints and $\mu(\varepsilon) q(\varepsilon)$ be the multiplier of the incentive compatible constraint.

$$\mathcal{L} = \max \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} W(\varepsilon) q(\varepsilon) d\varepsilon$$

$$+ \lambda_1 \left[F\left(K_1, \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} z_1(\varepsilon) q(\varepsilon) d\varepsilon\right) + (1 - \delta_k) K_1 - G_1 - \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} d_1(\varepsilon) q(\varepsilon) d\varepsilon - K_2 \right]$$

$$+ \lambda_2 \left[F\left(K_2, \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} z_2(\varepsilon) q(\varepsilon) d\varepsilon\right) + (1 - \delta_k) K_2 - G_2 - \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} d_2(\varepsilon) q(\varepsilon) d\varepsilon \right]$$

$$+ \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \mu(\varepsilon) q(\varepsilon) \left[\begin{aligned} & W(\varepsilon) - W(\underline{\varepsilon}) \\ & + \int_{\underline{\varepsilon}}^{\varepsilon} \left\{ \begin{aligned} & u'\big(d_1(\tilde{\varepsilon}) - e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon})\big) D_2 e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon}) \\ & - \beta v'\left(\frac{z_2(\tilde{\varepsilon})}{\Phi(h_1, e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon})) + \tilde{\varepsilon}}\right) \frac{z_2(\tilde{\varepsilon})\big[\Phi_e(h_1, e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon})) D_2 e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon}) + 1\big]}{\big[\Phi(h_1, e_1(\tilde{\varepsilon} \mid \tilde{\varepsilon})) + \tilde{\varepsilon}\big]^2} \end{aligned} \right\} q(\tilde{\varepsilon}) d\tilde{\varepsilon} \end{aligned} \right] d\varepsilon.$$

First-order conditions: Replacing $W(\varepsilon)$ by $u\big(d_1(\varepsilon) - e_1(\varepsilon \mid \varepsilon)\big) - v\left(\frac{z_1(\varepsilon)}{h_1}\right) + \beta\left[u\big(d_2(\varepsilon)\big) - v\left(\frac{z_2(\varepsilon)}{\Phi(h_1, e_1(\varepsilon \mid \varepsilon)) + \varepsilon}\right)\right]$

and then taking the first-order conditions gives

$$\big[d_1(\varepsilon)\big]: \quad (1 + \mu(\varepsilon)) u'\big(d_1(\varepsilon) - e_1(\varepsilon \mid \varepsilon)\big) + u''\big(d_1(\varepsilon) - e_1(\varepsilon \mid \varepsilon)\big) D_2 e_1(\varepsilon \mid \varepsilon) \int_{\varepsilon}^{\bar{\varepsilon}} \mu(\hat{\varepsilon}) q(\hat{\varepsilon}) d\hat{\varepsilon} = \lambda_1,$$

$$\big[d_2(\varepsilon)\big]: \quad (1 + \mu(\varepsilon)) \beta u'\big(d_2(\varepsilon)\big) = \lambda_2,$$

$$\big[K_2\big]: \quad \lambda_1 = \lambda_2 F_K\left(K_2, \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} z_2(\varepsilon) q(\varepsilon) d\varepsilon\right) + (1 - \delta_k),$$

$$\big[z_1(\varepsilon)\big]: \quad (1 + \mu(\varepsilon)) v'\left(\frac{z_1(\varepsilon)}{h_1}\right) \frac{1}{h_1} = \lambda_1 \left[F_Z\left(K_1, \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} z_1(\varepsilon) q(\varepsilon) d\varepsilon\right) \right],$$

$$[z_2(\varepsilon)]: \quad \frac{\beta(1+\mu(\varepsilon))}{\Phi(h_1,e_1(\varepsilon|\varepsilon))+\varepsilon}v'\left(\frac{z_2(\varepsilon)}{\Phi(h_1,e_1(\varepsilon|\varepsilon))+\varepsilon}\right)+\frac{\beta\Phi_e(h_1,e_1(\varepsilon|\varepsilon))D_2e_1(\varepsilon|\varepsilon)+1}{\left[\Phi(h_1,e_1(\varepsilon|\varepsilon))+\varepsilon\right]^2}\left[v'\left(\frac{z_2(\varepsilon)}{\Phi(h_1,e_1(\varepsilon|\varepsilon))+\varepsilon}\right)+\frac{z_2(\hat\varepsilon)v''\left(\frac{z_2(\varepsilon)}{\Phi(h_1,e_1(\varepsilon|\varepsilon))+\varepsilon}\right)}{\Phi(h_1,e_1(\varepsilon|\varepsilon))+\varepsilon}\right]\int_\varepsilon^{\bar\varepsilon}\mu(\hat\varepsilon)q(\hat\varepsilon)d\hat\varepsilon$$

$$= \lambda_2 F_Z\left(K_2, \int_{\underline\varepsilon}^{\bar\varepsilon} z_2(\varepsilon)q(\varepsilon)d\varepsilon\right),$$

Then, using these first-order conditions, we have the following equations:

$$u'\big(c_1(\varepsilon|\varepsilon)\big)-\beta u'\big(c_2(\varepsilon)\big)\big\{F_K(K_2,Z_2)+(1-\delta_k)\big\}=-\frac{\int_\varepsilon^{\bar\varepsilon}\mu(\hat\varepsilon)q(\hat\varepsilon)d\hat\varepsilon}{1+\mu(\varepsilon)}u''\big(c_1(\varepsilon|\varepsilon)\big)D_2e_1(\varepsilon|\varepsilon), \tag{C1a}$$

$$v'\left(\frac{z_1(\varepsilon)}{h_1}\right)\frac{1}{h_1}-u'\big(c_1(\varepsilon|\varepsilon)\big)F_Z(K_1,Z_1)=\frac{\int_\varepsilon^{\bar\varepsilon}\mu(\hat\varepsilon)q(\hat\varepsilon)d\hat\varepsilon}{1+\mu(\varepsilon)}u''\big(c_1(\varepsilon|\varepsilon)\big)D_2e_1(\varepsilon|\varepsilon)F_Z(K_1,Z_1), \tag{C1b}$$

$$v'\left(\frac{z_2(\varepsilon)}{h_2(\varepsilon|\varepsilon)}\right)\frac{1}{h_2(\varepsilon|\varepsilon)}-u'\big(c_2(\varepsilon)\big)F_Z(K_2,Z_2)=-\frac{\left[\Phi_e(h_1,e_1(\varepsilon|\varepsilon))D_2e_1(\varepsilon|\varepsilon)+1\right]}{(1+\mu(\varepsilon))\left[h_2(\varepsilon|\varepsilon)\right]^2}\left[v'\left(\frac{z_2(\varepsilon)}{h_2(\varepsilon|\varepsilon)}\right)+\frac{z_2(\varepsilon)}{h_2(\varepsilon|\varepsilon)}v''\left(\frac{z_2(\varepsilon)}{h_2(\varepsilon|\varepsilon)}\right)\right]\int_\varepsilon^{\bar\varepsilon}\mu(\hat\varepsilon)q(\hat\varepsilon)d\hat\varepsilon. \tag{C1c}$$

On the other hand, since

$$e_1(\hat\varepsilon|\varepsilon)\in\arg\max u\big(d_1(\hat\varepsilon)-e_1\big)-v\left(\frac{z_1(\hat\varepsilon)}{h_1}\right)+\beta\left[u\big(d_2(\hat\varepsilon)\big)-v\left(\frac{z_2(\hat\varepsilon)}{\Phi(h_1,e_1)+\varepsilon}\right)\right],$$

the value $e_1(\hat\varepsilon|\varepsilon)$ can be solved by the following equation

$$u'\big(d_1(\hat\varepsilon)-e_1\big)=\beta v'\left(\frac{z_2(\hat\varepsilon)}{\Phi(h_1,e_1)+\varepsilon}\right)\frac{z_2(\hat\varepsilon)\Phi_e(h_1,e_1)}{\left[\Phi(h_1,e_1)+\varepsilon\right]^2}. \tag{C2a}$$

Proposition 4. If Assumptions 1 and 4 hold, then the intertemporal wedges are positive, $\tau_{k_1}(\varepsilon)>0$ and the intratemporal wedges are negative in the first period, $\tau_{z_1}(\varepsilon)<0$, and positive in the second period, $\tau_{z_2}(\varepsilon)>0$, for any $\varepsilon\in(\underline\varepsilon,\bar\varepsilon)$.

Proof. According to Assumption 1, (C2a) can be written as

$$u'\big(d_1(\hat\varepsilon)-e_1(\hat\varepsilon|\varepsilon)\big)-\beta\kappa\gamma\frac{\left[z_2(\hat\varepsilon)\right]^\gamma\Phi_e\big(h_1,e_1(\hat\varepsilon|\varepsilon)\big)}{\left[\Phi\big(h_1,e_1(\hat\varepsilon|\varepsilon)\big)+\varepsilon\right]^{\gamma+1}}=0. \tag{C2b}$$

Taking the derivative on both sides of (C2b) with respect to ε, we get

$$\left\{u''\big(c_1(\hat\varepsilon|\varepsilon)\big)-\frac{\beta\kappa\gamma(\gamma+1)\left[z_2(\hat\varepsilon)\right]^\gamma}{\left[h_2(\hat\varepsilon|\varepsilon)\right]^{\gamma+2}}\left[\big(\Phi_e\big(h_1,e_1(\hat\varepsilon|\varepsilon)\big)\big)^2-h_2(\hat\varepsilon|\varepsilon)\Phi_{ee}\big(h_1,e_1(\hat\varepsilon|\varepsilon)\big)\right]\right\}D_2e_1(\hat\varepsilon|\varepsilon)=\frac{\beta\kappa\gamma(\gamma+1)\left[z_2(\hat\varepsilon)\right]^\gamma\Phi_e(h_1,e_1(\hat\varepsilon|\varepsilon))}{\left[h_2(\hat\varepsilon|\varepsilon)\right]^{\gamma+2}}$$

which leads to

$$D_2e_1(\hat\varepsilon|\varepsilon)=\frac{\frac{\beta\kappa\gamma(\gamma+1)\left[z_2(\hat\varepsilon)\right]^\gamma}{\left[h_2(\hat\varepsilon|\varepsilon)\right]^{\gamma+2}}\Phi_e\big(h_1,e_1(\hat\varepsilon|\varepsilon)\big)}{u''\big(c_1(\hat\varepsilon|\varepsilon)\big)-\frac{\beta\kappa\gamma(\gamma+1)\left[z_2(\hat\varepsilon)\right]^\gamma}{\left[h_2(\hat\varepsilon|\varepsilon)\right]^{\gamma+2}}\left[\big(\Phi_e\big(h_1,e_1(\hat\varepsilon|\varepsilon)\big)\big)^2-h_2(\hat\varepsilon|\varepsilon)\Phi_{ee}\big(h_1,e_1(\hat\varepsilon|\varepsilon)\big)\right]}. \tag{C2c}$$

Clearly, as $u''<0$, $\Phi_e>0$ and $\Phi_{ee}\leq0$, (C2c) gives $D_2e_1(\varepsilon|\varepsilon)<0$. Therefore, by (C1a) and $D_2e_1(\hat\varepsilon|\varepsilon)<0$, we get

$$u'\big(c_1(\varepsilon|\varepsilon)\big)<\beta u'\big(c_2(\varepsilon)\big)\big\{F_K(K_2,Z_2)+(1-\delta_k)\big\},$$

which implies that the capital wedge is positive, $\tau_{k_1}(\varepsilon) > 0$ for any $\varepsilon \in (\underline{\varepsilon}, \overline{\varepsilon})$.

Moreover, by (C1b) and $D_2 e_1(\varepsilon \mid \varepsilon) < 0$, we obtain

$$v'\left(\frac{z_1(\varepsilon)}{h_1}\right)\frac{1}{h_1} > u'\left(c_1(\varepsilon \mid \varepsilon)\right) F_Z\left(K_1, Z_1\right),$$

which implies that the labor wedge is negative in the first period, $\tau_{z_1}(\varepsilon) < 0$ for any $\varepsilon \in (\underline{\varepsilon}, \overline{\varepsilon})$.

On the other hand, using (C2c), we get

$$\Phi_e\left(h_1, e_1(\hat{\varepsilon} \mid \varepsilon)\right) D_2 e_1\left(\hat{\varepsilon} \mid \varepsilon\right) + 1 = \frac{u''\left(c_1\left(\hat{\varepsilon} \mid \varepsilon\right)\right) + \frac{\beta \kappa \gamma(\gamma+1)\left[z_2(\hat{\varepsilon})\right]^\gamma}{\left[h_2(\hat{\varepsilon}|\varepsilon)\right]^{\gamma+2}} h_2\left(\hat{\varepsilon} \mid \varepsilon\right)\Phi_{ee}\left(h_1, e_1\left(\hat{\varepsilon} \mid \varepsilon\right)\right)}{u''\left(c_1\left(\hat{\varepsilon} \mid \varepsilon\right)\right) - \frac{\beta \kappa \gamma(\gamma+1)\left[z_2(\hat{\varepsilon})\right]^\gamma}{\left[h_2(\hat{\varepsilon}|\varepsilon)\right]^{\gamma+2}}\left[\left(\Phi_e\left(h_1, e_1\left(\hat{\varepsilon} \mid \varepsilon\right)\right)\right)^2 - h_2\left(\hat{\varepsilon} \mid \varepsilon\right)\Phi_{ee}\left(h_1, e_1\left(\hat{\varepsilon} \mid \varepsilon\right)\right)\right]}.$$

Clearly, as $u'' < 0$, $\Phi_e > 0$ and $\Phi_{ee} \leq 0$, $\Phi_e\left(h_1, e_1(\varepsilon \mid \varepsilon)\right) D_2 e_1(\varepsilon \mid \varepsilon) + 1 > 0$. Therefore, by (C1c) and $\Phi_e\left(h_1, e_1(\varepsilon \mid \varepsilon)\right) D_2 e_1(\varepsilon \mid \varepsilon) + 1 > 0$, we obtain

$$v'\left(\frac{z_2(\varepsilon)}{h_2(\varepsilon|\varepsilon)}\right)\frac{1}{h_2(\varepsilon|\varepsilon)} < u'\left(c_2(\varepsilon)\right) F_Z\left(K_2, Z_2\right),$$

which implies that the labor wedge in the second period is positive, $\tau_{z_2}(\varepsilon) > 0$ for any $\varepsilon \in (\underline{\varepsilon}, \overline{\varepsilon})$. $\blacksquare$

Table 1. Exogenously calibrated parameters

Definition	Symbol	Value	Source/Note
Population			
Portion of the low type	π^L	0.5	By Assumption
Portion of the high type	π^H	0.5	By Assumption
Preference			
CRRA parameter	σ	2	Conesa et al. (2009)
Disutility of work scale	κ	1/3	Farhi and Werning (2013)
Disutility elasticity	γ	3	Farhi and Werning (2013)
Discount factor	β	0.542	0.96 annual ; Kapička and Neira (2015)
Final good production: Cobb-Douglas			
Productivity level	A	1	Normalization
Depreciation	δ_k	0.729	0.0833 annual; Conesa et al. (2009)
Capital share	ξ	0.36	Conesa et al (2009)
Human capital Technology: Ben-Porath			
Technology level	α	1	Normalization
Depreciation	δ_h	0.337	0.027 annual; Manuelli and Seshadri (2014)
Human capital share	η_1	0.486	Manuelli and Seshadri (2014)
Investment share	η_2	0.4	Manuelli and Seshadri (2014)
Initial human capital	h_1	1	Normalization
Initial advantage for type-	ε_1^L	0	Normalization
Tax system			
Capital income tax rate	τ_k	0.3	McDaniel (2007)
Labor income tax rate	τ_z	0.2	McDaniel (2007)
Government expenditure	G	0	By Assumption

Table 2. Endogenously matched parameters.

Calibrated parameter		Sim1	Sim2	Sim3	
Initial physical capital	K_1	0.081	0.081	0.081	
Initial advantage for H	ε_1^H	0.126	0.463	0.756	
Targeted data		**Sim1**	**Sim2**	**Sim3**	**Source**
Capital-output ratio	$K_1/F(K_1, K_1)$	0.1802	0.1815	0.1823	Peterman (2016)
Wage premium	z_3^H/z_3^L	1.2	1.8	2.4	Various sources

Table 3. Welfare gains over baseline

Economies	$\varepsilon_1^H=0.126$	$\varepsilon_1^H=0.463$	$\varepsilon_1^H=0.756$
Baseline	0%	0%	0%
Our model	0.871%	1.076%	1.427%
HCI observable	0.897%	1.121%	1.472%
HCI and type observable	0.916%	1.32%	1.76%

Note: Welfare gains are in terms of consumption equivalence. "Baseline" is the decentralized economy. "Our model" indicates the model that gives the constrained efficient allocation. "HCI observable" is our model except that human capital investment is observable. "HCI and type observable" is our model except that both human capital investment and agents' types are observable.
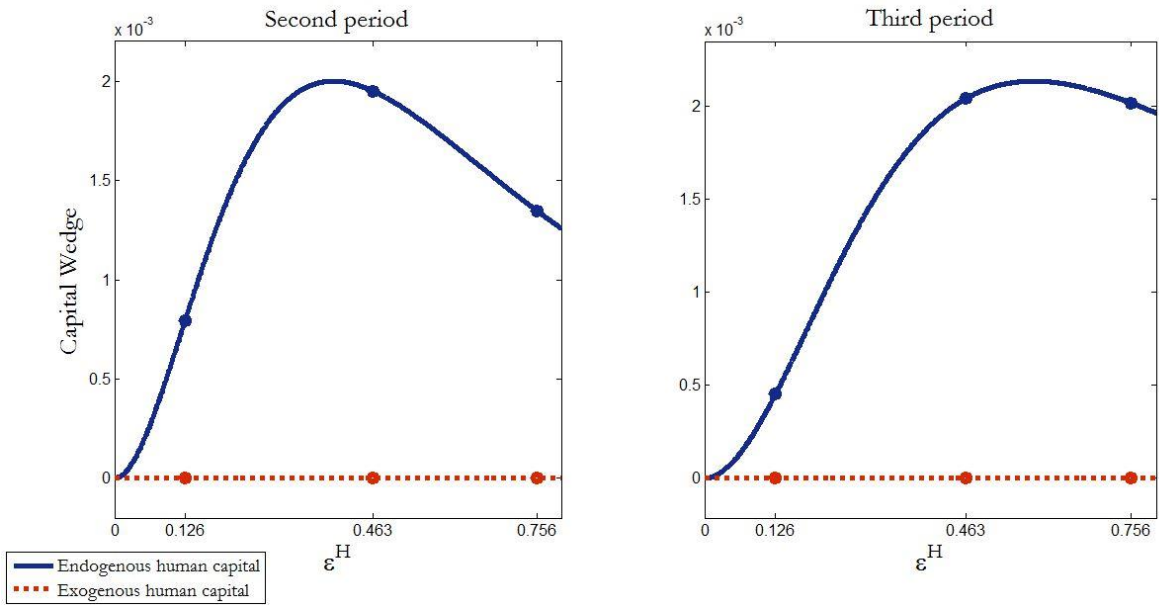
Figure 1. Capital wedges for the low type under different initial advantages of human capital (models with vs. without endogenous human capital).
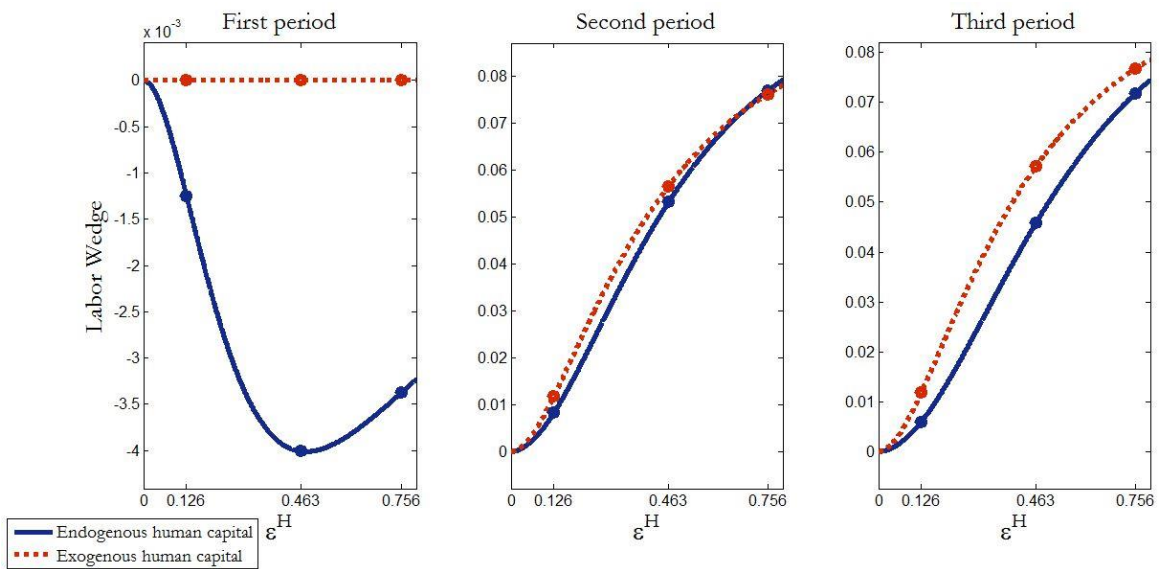

Figure 2. Labor wedges for the low type under different initial advantages of human capital (models with vs. without endogenous human capital).
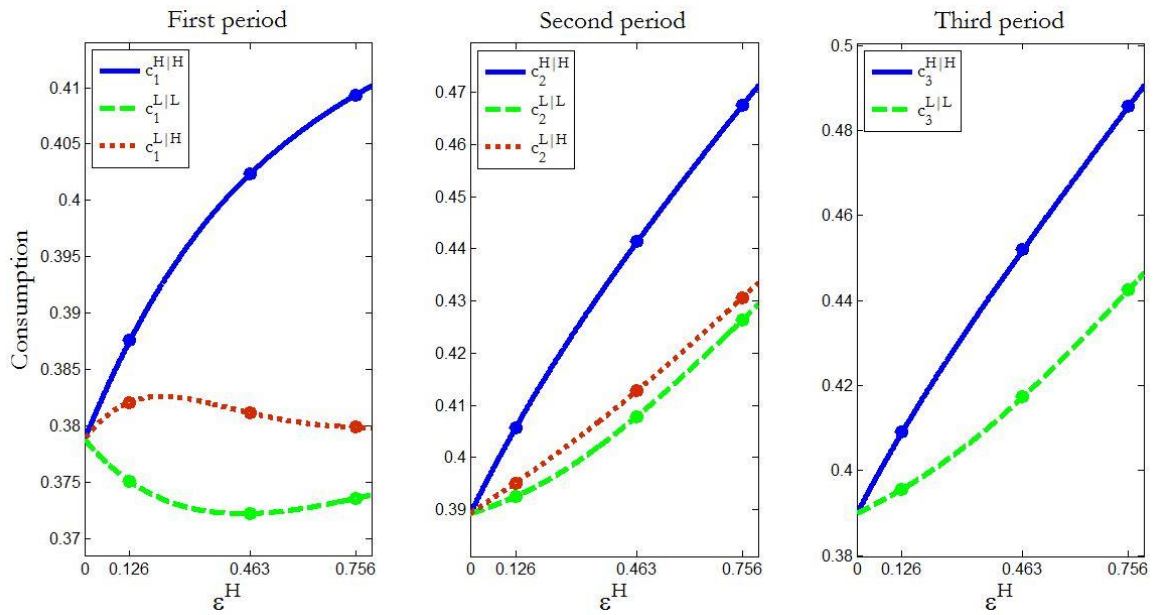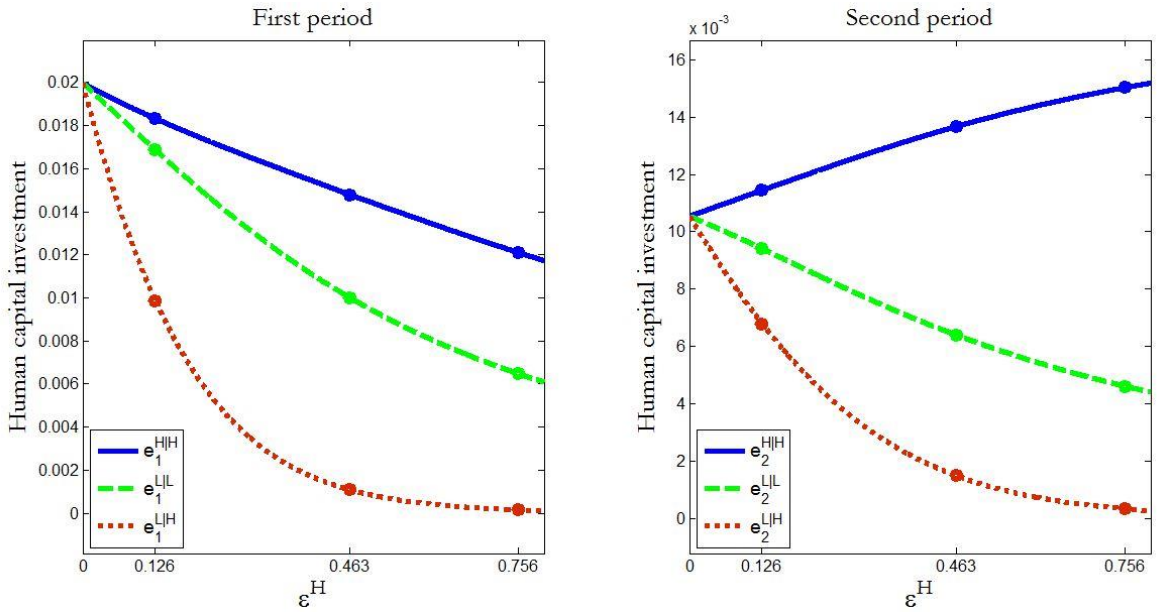
Figure 3. Consumption for different strategies.

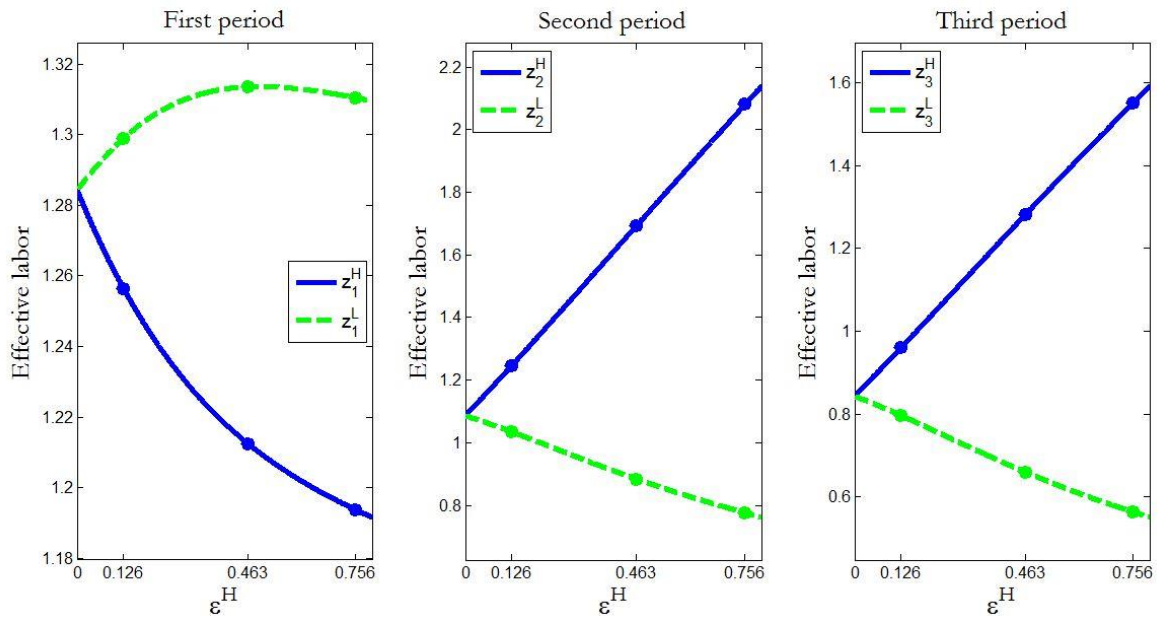**Figure 4. Human capital investment for different strategies.**
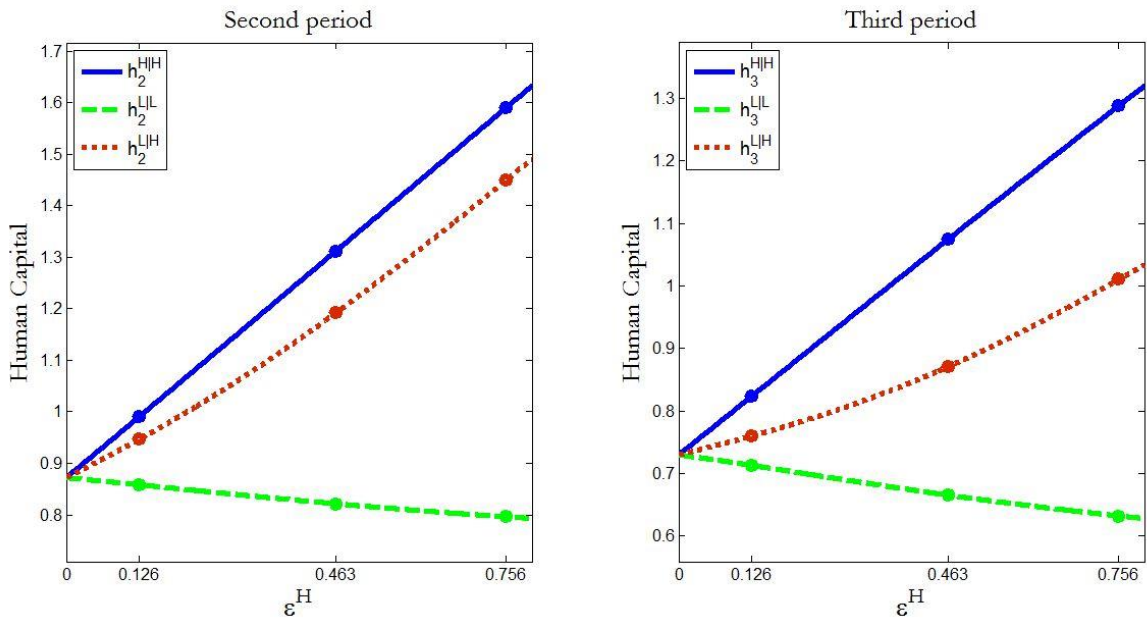
Figure 5. Effective labor for different strategies.

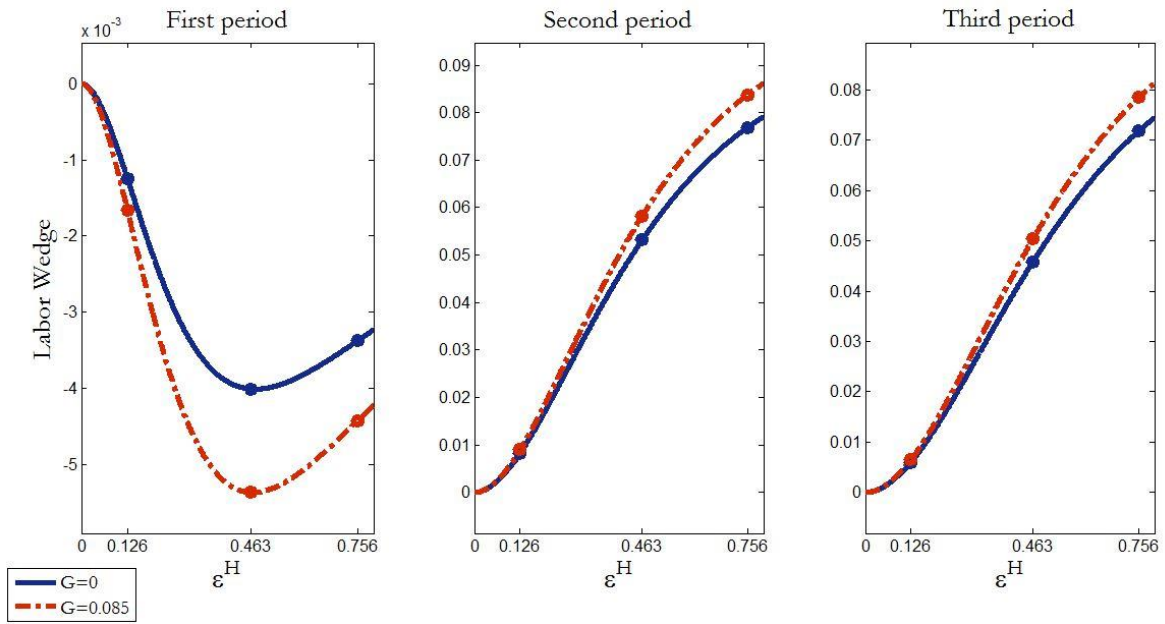**Figure 6. Human capital level for different strategies.**

Figure 7. Labor wedges for different government expenditure.

Appendix Table

$$\nabla^2 \mathcal{L} = \begin{bmatrix}
u''(c_0) & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
0 & \ddots & 0 & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\vdots & 0 & \beta^T u''(c_T) & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
0 & \cdots & 0 & -v''(l_0) & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
0 & \cdots & 0 & 0 & -\beta v''(l_1) & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_1 w_1^*\left(1-\tau_{z_1}^{r*}\right)-\eta_1 & 0 & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & 0 & \ddots & 0 & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & 0 & \ddots & 0 & \vdots \\
0 & \cdots & 0 & 0 & 0 & 0 & -\beta^{T-1}v''(l_{T-1}) & 0 & 0 & 0 & 0 & \cdots & 0 & \vdots & 0 & \lambda_{T-1}w_{T-1}^*\left(1-\tau_{z_{T-1}}^{r*}\right)-\eta_{T-1} & 0 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & -\beta^T v''(l_T) & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \lambda_T w_T^*\left(1-\tau_{z_T}^{r*}\right)-\eta_T \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \psi_{ee}(h_0,e_0) & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \frac{\partial^2}{\partial e^2}\psi_1^i(h_1,e_1) & 0 & \vdots & \frac{\partial^2}{\partial h \partial e}\psi_1^i(h_1,e_1) & 0 & \vdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & 0 & \ddots & 0 & 0 & \ddots & 0 & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \frac{\partial^2}{\partial e^2}\psi_{T-1}^i(h_{T-1},e_{T-1}) & 0 & 0 & \frac{\partial^2}{\partial h \partial e}\psi_{T-1}^i(h_{T-1},e_{T-1}) & 0 \\
0 & \cdots & 0 & 0 & \lambda_1 w_1^*\left(1-\tau_{z_1}^{r*}\right)-\eta_1 & 0 & \cdots & 0 & 0 & 0 & \frac{\partial^2}{\partial e \partial h}\psi_1^i(h_1,e_1) & 0 & 0 & \frac{\partial^2}{\partial h^2}\psi_1^i(h_1,e_1) & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & 0 & \ddots & \vdots & \vdots & \vdots & 0 & \ddots & 0 & 0 & \ddots & 0 & \vdots \\
0 & \cdots & 0 & 0 & \vdots & 0 & \lambda_{T-1}w_{T-1}^*\left(1-\tau_{z_{T-1}}^{r*}\right)-\eta_{T-1} & 0 & 0 & 0 & \vdots & 0 & \frac{\partial^2}{\partial e \partial h}\psi_{T-1}^i(h_{T-1},e_{T-1}) & 0 & 0 & \frac{\partial^2}{\partial h^2}\psi_{T-1}^i(h_{T-1},e_{T-1}) & 0 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \lambda_T w_T^*\left(1-\tau_{z_T}^{r*}\right)-\eta_T & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\partial^2}{\partial h^2}\psi_T^i(h_T,e_T)
\end{bmatrix}$$

Appendix matrix.