UNSW
THE UNIVERSITY OF NEW SOUTH WALES
SYDNEY • AUSTRALIA

University of New South Wales
School of Economics

Honours Thesis

Empty Nest: Optimal Dynamic Taxation with College Savings

Author:
Gary Wang
Student ID: 5075803

Supervisors:
Gonzalo Castex
Pei-Cheng Yu

Bachelors of Economics (Economics) (Honours)
and
Bachelors of Commerce (Finance, Business Strategy and Economic Management)

22nd November, 2019

Acknowledgements

To begin, I would like to thank my supervisors Gonzalo Castex and Pei-Cheng Yu for their endless support, and the tremendous amount of time spent on providing me with feedback and ideas to work with. I appreciate the patience you both had with me whenever I was stuck or confused, and overall I had an amazing experience with writing this thesis because of you guys. In particular, thank you PC for the idea of this thesis. Public finance is quite new to me and I appreciate the encouragement to give it a go.

I would like to also thank the academics and staff in the economics faculty, particularly Stanley Cho, Pratiti Chatterjee, Gabriele Gratton, Anuj Bhowmik, Pauline Grosjean, Federico Masera, Fanghua Li, Tess Stafford and Sarah Walker for expanding my knowledge in economics, providing me with feedback on my thesis, and making the Honours Program awesome.

To the Honours Cohort of 2019, you guys have been such an amazing bunch of people. The teamwork and selfless assistance throughout the year have made the journey bearable (particularly during term 1's micro and macro fiasco). In particular, I would like to thank the following: Carol for being my assignment buddy in term 1 but then ditching me in Policy Evaluation, for failing to keep up with our netflix shows, and also for being a semi-nice person when you choose to be (don't cry) ; Rhea for sharing Gonzalo and PC with me, and sorry for stealing PC for most of Term 1 ; Robert for carrying me in Policy Evaluation, you are a legend ; Jonathan for making me add people to the mailing list every week, and including me in your killer examples.

Finally, to my family and friends, thank you for your support. Whether it be tagging me in memes about my thesis, asking how my thesis is going, enticing me into playing games or going out when I shouldn't be, etc. Even though I know that none of you will read this, I appreciate you all.

Contents

Chapter 1
Introduction

Despite the large financial burden attached to college and university education, post-secondary education is common amongst young individuals looking to increase their human capital and improve their employability in the labour market. According to the National Center for Education Statistics, around 19.9 million students in the U.S. will be attending a college or university by the end of 2019. Unsurprisingly, parents and family members have access to tax advantaged college funds such as the 529 plan in the U.S. and the Registered Education Savings Program (RESP) in Canada. These saving accounts allow parents and family to begin saving early to lessen the financial burden and even encourage a college education for their children. However, this investment is risky because the child may not ultimately attend college in the future.

The aim of this paper is to follow the Mirrlees taxation framework to characterise the optimal college savings account when parents faces future uncertainties on their child's outcome. In particular, this paper considers two models; an intergenerational model where an altruistic parent works and saves towards their child's college education, and a human capital life-cycle model to capture the potential benefits of the college investment for the child. The implicit assumption is that investment is made because of the parent's inherent care for their child's future wellbeing, and that the child agrees with the decisions made by the parent.

Like any dynamic Mirrlees taxation framework, agents in this environment differ in productivity. In particular, the child's productivity is determined by his or her human capital, consisting of their own innate ability and any education investment. The government in this framework is interested in the outcomes of the child, and thus ensures that the welfare of the child is taken into account. The government then implements policies by designing the optimal allocations for both the parent and the child. However, productivity and innate ability are both unobservable by the government, which causes informational friction whenever the government wishes to redistribute. The mechanism design approach is used to solve the government's problem. This method allow the underlying frictions within the model determine the optimal taxes and policies that achieves the optimal allocations. In this approach, dynamic incentive compatibility is ensured such that agents self-select the optimal allocation. The government can then implement these allocations through the optimal policies.

The optimal intergenerational wedge in this paper characterises the optimal design for a college savings account. This wedge is similar to the one found in Farhi and Werning (2007), but their environment does not consider an overlap between the parent and the child's life-cycle, which is captured in this paper. The theoretical results find that whenever the government places any direct weighting on the child's welfare, then subsidies to college savings should be made even before the parent realises whether or not their child should attend college. Furthermore, the optimal subsidy is decreasing in the parent's wealth and increasing in the relative weighting of the government versus the parent's care for the child's future.

In addition, I look at the optimal intertemporal and intratemporal conditions in this model. The intertemporal conditions allow for the characterisation of the optimal personal savings distortions. Results match the general inverse Euler condition for both the parent and the child, implying a positive optimal tax on personal savings for both agents. Next, the optimal intratemporal distortions provides an understanding of how the marginal tax rate affect labour supply decisions. In this intergenerational model, distortions to the labour wedge consists of intratemporal and intertemporal components that are similar to Golosov, Troshkin, and Tsyvinski (2016), suggesting that the government can take advantage of distortions between periods to achieve the optimal allocations. Since there is an overlap between the parent and child in the model, the trade-offs between the consumption for the parent and child are captured in the parent's labour wedge.

Related Literature. This paper relates to existing literature that followed on from the Mirrlees (1971) framework on optimal taxation by considering a dynamic environment and intergenerational human capital investment. Human capital investment is costly and may not necessarily guarantee positive returns. Lochner and Monge-Naranjo (2014) discusses rising labour market uncertainties and trends such as major increases in the real cost of obtaining higher education over the past decade[1]. Given this large financial burden, many college students require financial aid in the form of student loans and family contributions. A study by Gale and Scholz (1994) finds that inter vivos contributions towards education are sizeable but does not affect college enrolment significantly. However, more recent studies such as Dynarski (2003) and Belley and Lochner (2007) finds that family income in the U.S. has played a more significant role in college attendance and completion during the early 2000s.

Parental transfers are an important source of financial aid for students and are considered in many of the literature on education policies. Castex (2017) looks

[1]Lochner and Monge-Naranjo (2014) found that the average net cost of attendance at public four-year colleges increased by 64% between 1990 and 2012.

at how various benefits and costs of college affect college enrolments when human capital is risky. According to his model, parental transfers can increase enrolment by as much as 14 percentage points[2] amongst high-ability individuals. Abbott, Gallipoli, Meghir, and Violante (2018) examines policies aimed at alternative financial aid including intergenerational transfers and acknowledges the idea that both cognitive and non-cognitive skills of a child depends on their parent's cognitive skills and education level, which is captured in my model.

Since this paper considers optimal taxation with human capital investment, the determination of a child's future productivity is similar to Findeisen and Sachs (2016)'s study on the optimal design of education finance and tax systems. In their environment, agents differ in initial innate ability, and make the risky decision on whether to invest into education before entering into the labour market, and this decision is observable by the government. Grochulski and Piskorski (2010) also explores a similar environment where initial human capital investment affects future workforce ability but does not consider differences in innate ability. Other literature on human capital investment and optimal taxation also considers human capital acquisition during the working periods such as in Kapicka and Neira (2015), as well as human capital accumulation over a life-cycle in Stantcheva (2015b). However, this paper focuses solely on human capital acquisition through a college education.

This paper also relates to literature on optimal taxation with inter vivos transfers. Farhi and Werning (2010) explores optimal taxation for bequests by considering an environment with altruistic parents and heterogeneous productivity. They find that estate tax should be progressive and that parents should face a marginal subsidy on bequests, and that these results could also apply to educational investments. However, their model does not consider human capital accumulation for the child. Stantcheva (2015a) considers the interplay between the transfer of resources through education versus bequests. Similarly, Pavoni and Yazici (2016) examine how intergenerational disagreements could affect the optimal taxation of bequests and inter vivos transfers. However, my work focuses on characterising the optimal savings account rather than the optimisation of joint decisions on bequests and inter vivos education transfers.

This paper contributes to current literature on optimal taxation and human capital investment by characterising the optimal taxation policies on inter vivos education transfers. It incorporates a human capital life-cycle model in an intergenerational framework to provide a simple understanding of the optimal design for a college savings account.

The structure of this paper is as follows. Section 2 contains the structure of the model

[2]From the model in Castex (2017), parental transfer was set to zero and I compare the enrolment rate of 4-year colleges for the highest quartile ability individuals in high-income households.

and the approach to solving the government's problem. Section 3 defines the three optimal wedges that are considered in this paper and examines the implications of the theoretic results that are found by solving the government's problem. Section 4 looks at an implementation of the optimal policies based on observable information such education investment and income, and comparing this to existing policies. Finally, Section 5 discusses potential research in the future, and concludes the paper.

Chapter 2
Model Setup

The intergenerational model consists of parents and children denoted as p and k respectively, with heterogeneous productivity $\vartheta_i \in \Theta = \left[\underline{\vartheta}, \overline{\vartheta}\right]$, for $i \in \{p, k\}$. Each parent has one child. Agents, either parent or child, of productivity type ϑ_i supply l_t^i unit of labour to produce output $y_t^i = \vartheta_i l_t^i$ in period t. The disutility of supplying l_t^i labour is h. Agent i derives utility u from consuming c_t^i in period t. The agent's productivity and labour supply are private information, but the government can observe output.

The model has three periods: $t = 0, 1, 2$. The parent is altruistic and lives in $t = 0, 1$. The life-cycle of the parent can be summarised in Figure 1.

In $t = 0$, the parent realises their own productivity type ϑ_p drawn from a differentiable probability distribution function (pdf) $f^p(\vartheta_p)$ with a cumulative distribution function (cdf) $F^p(\vartheta_p)$[1]. The net flow utility at $t = 0$ is:

$$u\left(c_0^p(\vartheta_p)\right) - h\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)$$

Where $u'(.) > 0$, $u''(.) < 0$, $h'(.) > 0$, $h''(.) > 0$.

In $t = 1$, the parent receives a private signal $\sigma \in \{L, H\}$ on their child's innate ability where L and H denote low and high innate ability respectively. The probability distribution of innate ability depends on the parent's productivity. More specifically, the probability distribution $\Pr\left(H|\vartheta_p\right) = g\left(\vartheta_p\right)$ is strictly increasing with ϑ_p. The child's signal may be thought of as early signs observed by the parent during childhood. Parents who are more productive tend to have better parental investments that positively affects their child's ability (Anger and Heineck, 2010).

Finally, the parent also chooses whether to invest in their child's college education $e \in \{0, 1\}$, where education is a binary choice on whether to send the child to college $(e = 1)$ or not $(e = 0)$, and this decision is observable by the government.

[1] I do not place any restrictions on this distribution yet.

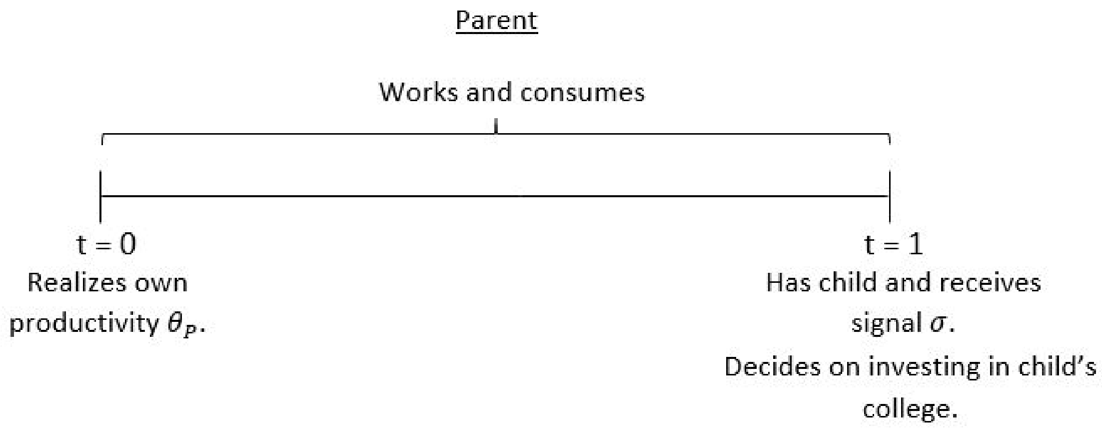

Figure 2.1: Parent's Lifecycle

The utility flow at $t = 1$ is denoted as $U_1^p(c_1^p, y_1^p; \sigma, \vartheta_p, e)$ if the child is σ type. More specifically:

$$U_1^p(c_1^p, y_1^p; \sigma, \vartheta_p, e) = u\left(c_1^p(\sigma, \vartheta_p)\right) - h\left(\frac{y_1^p(\sigma, \vartheta_p)}{\vartheta_p}\right) + \alpha U_1^k\left(c_t^k, y_2^k; \sigma, e\right)$$

Where $\alpha > 0$ is the altruism parameter and U_1^k is the child's discounted life-time utility.

Given that $\beta_1^p < 1$ is the discounting factor for the parent at $t = 1$, the discounted life-time utility for the parent is then:

$$U_0^p(c_0^p, y_0^p; \vartheta_p) = u(c_0^p(\vartheta_p)) - h\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right) + \beta_1^p \mathbb{E}_\sigma\left[U_1^p(c_1^p, y_1^p; \sigma, \vartheta_p, e)\right]$$

The child lives in $t = 1, 2$. Figure 2 summarises the life-cycle timeline for the child. At $t = 1$, the child consumes and goes to college or not depending on the parent's decision. Here, I assume that there are no disagreements between the parent and child's intertemporal trade-offs[2]. In $t = 2$, the child begins to work and productivity ϑ_k is realised.

[2]This assumption is fair given that college saving plans can sometimes be used for qualified education expenses outside of college education. Hence there are flexibility in this assumption

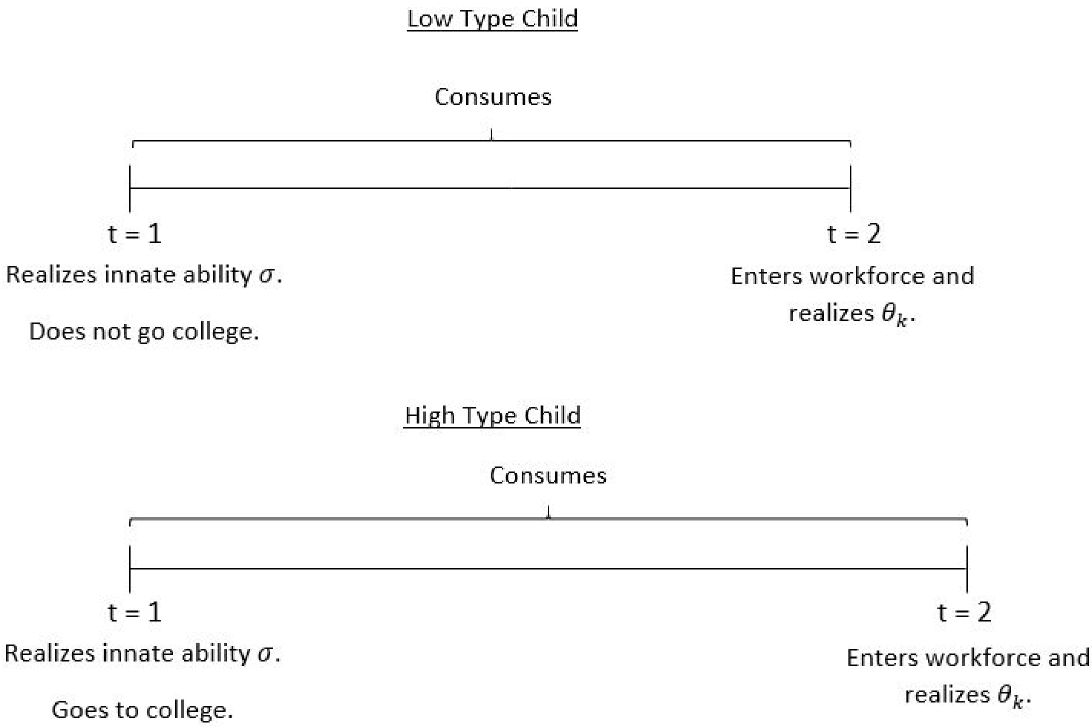

Figure 2.2: Low and High Type Child's Lifecycle

The child's productivity depends on his or her human capital $\varkappa(\sigma, e)$, where $\varkappa$ is made up of the child's innate ability and education investment. More specifically, ϑ_k is drawn from a continuous pdf, $f^k\left(\vartheta_k|\varkappa\left(\sigma, e\right)\right)$, with cdf $F^k\left(\vartheta_k|\varkappa\left(\sigma, e\right)\right)$ which satisfies first order stochastic dominance: $F^k\left(\vartheta_k|\varkappa\right) < F^k\left(\vartheta_k|\varkappa'\right)$ for all ϑ_k where $\varkappa > \varkappa'$. Intuitively, this means that children with higher human capital are more likely to have a higher productivity type. This characterisation captures the risk of human capital investment, even though on average the expected return to college education is positive (Castex, 2017; Cunha and Heckman, 2007).

The discounted life-time utility of a (σ, ϑ_p) child at $t = 1$ when parents report type σ' is:

$$U_1^k\left(c_1^k; \sigma', e, \sigma\right) = \beta_1^e u\left(c_1^k(\sigma', \vartheta_p)\right) + \beta_2^e \int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[U_2^k(c_2^k, y_2^k; \sigma', e)\right] f^k\left(\vartheta_k|\varkappa\left(\sigma, e\right)\right) d\vartheta_k.$$

where:

$$U_2^k(c_2^k, y_2^k; \sigma', e) = u\left(c_2^k(\sigma', \vartheta_p, \vartheta_k)\right) - h\left(\frac{y_2^k(\sigma', \vartheta_p, \vartheta_k)}{\vartheta_k}\right)$$

Since innate ability is unobservable, the reported σ' can be different to the true σ[3].

[3] I make this distinction so that it is easier to define the incentive compatibility constraint at $t = 1$ in the government's problem.

Furthermore, $\beta_t^e < 1$ is the discount factor for the child at $t = 1, 2$ for $e \in \{0, 1\}$ to account for the different period lengths of the pre-working and working period for a college child versus the non-college child since $t = 1$ is longer for those who undertake college education.

2.0.1 The Planning Problem

The government's problem is solved through a mechanism design approach whereby the government implements policies to achieve the optimal allocations over consumption and output, given the information frictions between the government and the agents. The direct mechanism consists of designing both the parent's allocations and the child's allocations, and is given as:

$$P = \{c_0^p(\vartheta_p), y_0^p(\vartheta_p), c_1^p(\sigma, \vartheta_p), y_1^p(\sigma, \vartheta_p)\}$$

$$K = \{c_1^k(\sigma, \vartheta_p), c_2^k(\sigma, \vartheta_p, \vartheta_k), y_2^k(\sigma, \vartheta_p, \vartheta_k)\}$$

The revelation principle for this direct mechanism approach requires (P, K) to be incentive compatible in each period $t = 0, 1, 2$, so that agents do not have an incentive to misreport information that is unobservable by the planner. Since this is a sequential mechanism problem, the dynamic incentive constraints are ensured backwards.

The ex-post incentive compatibility constraint at $t = 2$ ensures that the child will report their own type ϑ_k truthfully. Given any ϑ_p and σ report from $t = 0, 1$, the incentive constraint at $t = 2$ is:

$$u(c_2^k(\sigma, \vartheta_p, \vartheta_k)) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\vartheta_k}\right) \geq u(c_2^k(\sigma, \vartheta_p, \vartheta_k')) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta_k')}{\vartheta_k}\right)$$

Where ϑ_k is the true productivity of the child and $\vartheta_k' \neq \vartheta_k$.

The incentive constraint at $t = 1$ ensures that for all ϑ_p reports at $t = 1$, the parent correctly reports σ. The focus is on a separating mechanism where the parent will always send the child to college if and only if the child is observed to be $\sigma = H$[4].

If the child is observed $\sigma = H$:

$$u(c_1^p(H, \vartheta_p)) - h\left(\frac{y_1^p(H, \vartheta_p)}{\vartheta_p}\right) + \alpha U_1^k(H, e = 1) \geq u(c_1^p(L, \vartheta_p)) - h\left(\frac{y_1^p(L, \vartheta_p)}{\vartheta_p}\right) + \alpha U_1^k(H, e = 0)$$

[4]This means that if the parent does not send their child to college despite observing $\sigma = H$, then the parent reports $\sigma' = L$ and consumes as if the child is a L type.

If the child is observed $\sigma = L$:

$$u(c_1^p(L,\vartheta_p)) - h\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right) + \alpha U_1^k(L, e=0) \geq u(c_1^p(H,\vartheta_p)) - h\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right) + \alpha U_1^k(L, e=1)$$

The ex-ante incentive constraint at $t = 0$ ensures that the parent truthfully reports their own ϑ_p, such that:

$$u(c_0^p(\vartheta_p)) - h\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right) + \beta_1^p \mathbb{E}_\sigma\left[U_1^p(\vartheta_p;\sigma,\vartheta_p)\right] \geq u(c_0^p(\vartheta_p')) - h\left(\frac{y_0^p(\vartheta_p')}{\vartheta_p}\right) + \beta_1^p \mathbb{E}_\sigma\left[U_1^p(\vartheta_p';\sigma,\vartheta_p)\right]$$

Where ϑ_p is the true productivity of the parent and $\vartheta_p \neq \vartheta_p'$.

The planning problem considers the welfare of the parents and the children, with welfare weight γ given to the parents, and $(1-\gamma)$ to the children. Furthermore, the assumption is made such that the planner discounts period $t = 1$ at the same rate as the parent. Then, the constrained planning problem maximises:

$$\gamma \int_{\underline{\vartheta}}^{\overline{\vartheta}} U_0^p(\vartheta_p) dF(\vartheta_p) + (1-\gamma)\beta_1^p \left[\int_{\underline{\vartheta}}^{\overline{\vartheta}} g(\vartheta_p) U_1^k(H,1) dF(\vartheta_p) + \int_{\underline{\vartheta}}^{\overline{\vartheta}} (1-g(\vartheta_p))\, U_1^k(L,0) dF(\vartheta_p)\right]$$

Subject to the incentive constraints and the feasibility constraint.

The feasibility constraint is simply the resource constraint such that the total output is equal to the total consumption by the parents and children across the economy plus any education resources e_σ in period $t = 1$.
The resource constraint can be expressed in present-value terms with the rates of return R_1^p , R_1^e , $R_2^e > 0$ for period $t = 1, 2$ and $e = 0, 1$, such that:

$$\int_{\underline{\vartheta}}^{\overline{\vartheta}} \left\{ Y_0 + \frac{g(\vartheta_p)}{R_1^p}\left[Y_1^1 + \frac{1}{R_2^k(1)}\int_{\underline{\vartheta}}^{\overline{\vartheta}} Y_2^1 dF^k(\vartheta_k|H,1)\right] + \frac{1-g(\vartheta_p)}{R_1^p}\left[Y_1^0 + \frac{1}{R_2^k(0)}\int_{\underline{\vartheta}}^{\overline{\vartheta}} Y_2^0 dF^k(\vartheta_k|L,0)\right]\right\} d\vartheta_p \geq 0$$

where:

$$\begin{aligned} Y_0 &= y_0^p(\vartheta_p) - c_0^p(\vartheta_p) \\ Y_1^e &= y_1^p(\sigma,\vartheta_p) - c_1^p(\sigma,\vartheta_p) - \frac{c_1^k(\sigma,\vartheta_p) + e_\sigma}{R_1^k(e)} \\ Y_2^e &= y_2^k(\sigma,\vartheta_p,\vartheta_k) - c_2^k(\sigma,\vartheta_p,\vartheta_k) \end{aligned}$$

For simplicity, the assumption is made such that $\beta_1^p R_1^p = \beta_1^e R_1^k(e) = \beta_2^e R_2^k(e) = 1$.

2.0.2 Characterising Incentive Compatibility

Before solving the constrained optimisation problem, the incentive constraints can be simplified. The ex-post compatibility constraint can be simplified as follows:

Lemma 1 *For any* σ *and* ϑ_p, K *is ex-post incentive compatible if and only if* ***(i)*** $y_2^k(\sigma, \vartheta_p, \vartheta_k)$ *is non-decreasing in* ϑ_k *and* ***(ii)*** $U_2^k(\sigma, \vartheta_p, \vartheta_k)$ *is absolutely continuous in* ϑ_k*, with* $\frac{\delta U_2^k(\sigma,\vartheta_p,\vartheta_k)}{\delta\vartheta_k} = \frac{y_2(\sigma,\vartheta_p,\vartheta_k)}{\vartheta_k^2} h' \left(\frac{y_2(\sigma,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)$.

Lemma 1 characterises the general conditions for K to be ex-post incentive compatible[5]. As a result, the child's utility at $t = 2$ can be rewritten as:

$$U_2^k(\sigma, \vartheta_p, \vartheta_k) = U_2^k(\sigma, \vartheta_p, \underline{\vartheta}) + \int_{\underline{\vartheta}}^{\vartheta_k} \frac{y_2(\sigma, \vartheta_p, k)}{k^2} h' \left(\frac{y_2(\sigma, \vartheta_p, k)}{k}\right) dk$$

Next, the following Lemma characterises the ex-ante incentive constraint at $t = 1$.

Lemma 2 *Suppose* K *is ex-post incentive compatible. Then if* ***(i)*** *The returns to education for high innate ability is greater than for low innate ability agents such that:*

$$\beta_2^{e_1} |F(\vartheta_k|\varkappa(H, e = 1) - F(\vartheta_k|\varkappa(L, e = 1))| - \beta_2^{e_0} |F(\vartheta_k|\varkappa(H, e = 0) - F(\vartheta_k|\varkappa(L, e = 0))| \geq 0$$

and ***(ii)*** *Output is increasing in innate ability for all* ϑ_p, ϑ_k*:* $y_2(H, \vartheta_p, \vartheta_k) \geq y_2(L, \vartheta_p, \vartheta_k)$*, then* $U_1^p(H, e = 1; \vartheta_p) = U_1^p(H, e = 0; \vartheta_p)$ *at the optimum.*

This tells us that as long as Lemma 1 holds and that the conditions for (i) and (ii) of Lemma 2 are true, then the only deviation to worry about is for the parent of a high innate ability child to misreport their child as low innate ability.
The incentive compatibility constraint at $t = 0$ can be simplified in the same way as the ex-post incentive constraint from Lemma 1.

Lemma 3 *Suppose that K is ex-post incentive compatible and that* $U_1^p(H, e = 1; \vartheta_p) = U_1^p(L, e = 0; \vartheta_p)$ *at the optimum. Then (P,K) is ex-ante incentive compatible if and only if* ***(i)*** $y_0^p(\vartheta_p)$ *is non-decreasing in* ϑ_p *and* ***(ii)*** $U_0^p(\vartheta_p)$ *is absolutely continuous in* ϑ_p *with:*

$$\frac{\delta U_1^p(\vartheta_p)}{\delta\vartheta_p} = \frac{y_0^p(\vartheta_p)}{\vartheta_p^2} h' \left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right) + \beta_1^p g(\vartheta_p) \left[\frac{y_1^p(H, \vartheta_p)}{\vartheta_p^2} h' \left(\frac{y_1^p(H, \vartheta_p)}{\vartheta_p}\right)\right]$$
$$+\beta_1^p (1 - g(\vartheta_p)) \left[\frac{y_1^p(L, \vartheta_p)}{\vartheta_p^2} h' \left(\frac{y_1^p(L, \vartheta_p)}{\vartheta_p}\right)\right] + \alpha\beta_1^p g'(\vartheta_p) \left[U_1^k(H, e = 0) - U_1^k(L, e = 0)\right]$$

[5] This is also known as the revenue equivalence.

Lemma 3 characterises the general conditions for (P, K) to be ex-ante incentive compatible at $\mathrm{t} = 0$.

Chapter 3
Theoretical Results - Optimal Wedges

Solving the planning problem[1] allows for the characterisation of the intertemporal and intratemporal conditions found in this environment. These conditions then determine the optimal tax mechanism. Specifically, this section provides an understanding of the optimal distortions to the savings and labour wedges for an altruistic parent that decides on making a risky investment towards their child. From these results, the government can then design an optimal college savings account.

3.0.1 Intertemporal Wedges

The intertemporal wedge characterises the distortions from redistribution on consumption and saving decisions between two periods when there is asymmetric information. This section focuses on the agent's savings for future consumption. Given that $\beta_1^p R_1^p = \beta_1^e R_1^k(e) = \beta_2^e R_2^k(e) = 1$:
The intertemporal wedge for a parent of type ϑ_p at $t = 0$ is defined by:

$$\tau_0^c(\vartheta_p) = 1 - \frac{u'(c_0(\vartheta_p))}{\mathbb{E}_\sigma\left[u'(c_1(\sigma, \vartheta_p))\right]}$$

The intertemporal wedge for a (ϑ_p, σ) type child at $t = 1$ is defined by:

$$\tau_1^c(\sigma, \vartheta_p) = 1 - \frac{u'(c_1^k(\sigma, \vartheta_p))}{\mathbb{E}_{\vartheta_k}\left[u'(c_2^k(\sigma, \vartheta_p, \vartheta_k))\right]}$$

If $\tau^c = 0$, then the intertemporal Euler holds such that agents are consuming efficiently across periods by equating their marginal rate of substitution for consumption, with the price ratio of consumption between the two periods. When $\tau^c > 0$, the agent is consuming too much and saving lower than the efficient amount. If $\tau^c < 0$, the parent is saving more than the efficient amount.

[1]The constrained problem was solved by setting up a Hamiltonian. The technique and first order conditions can be found in A.0.2.

Proposition 1 *The optimal allocation satisfies the inverse Euler equations for both agents:*

$$\frac{1}{u'(c_0^p(\vartheta_p))} = \mathbb{E}_\sigma\left[\frac{1}{u'(c_1^p(\sigma,\vartheta_p))}\right] \tag{1}$$

$$\frac{1}{u'(c_1^k(\sigma,\vartheta_p))} = \mathbb{E}_{\vartheta_k}\left[\frac{1}{u'(c_2^k(\sigma,\vartheta_p,\vartheta_k))}\right] \tag{2}$$

Proposition 1[2] characterises the optimal intertemporal wedge for the parent and child when the government wishes to redistribute but exist in an environment where productivity and innate ability are heterogeneous and unobservable. The implications of the inverse Euler equation can be understood by considering equation (1). Jensen's inequality states that $\mathbb{E}\left[\frac{1}{x}\right] > \frac{1}{\mathbb{E}[x]}$ if $\frac{1}{x}$ is convex, meaning that $\mathbb{E}_\sigma\left[\frac{1}{u'(c_1^p(\sigma,\vartheta_p))}\right] > \frac{1}{\mathbb{E}_\sigma[u'(c_1(\sigma,\vartheta_p))]}$. This implies that a positive marginal savings distortion $\tau_0^c(\vartheta_p) > 0$ for any ϑ_p is optimal to achieve the constrained efficient allocations.

The result here may seem counter-intuitive given that the parent needs to save up for their child. However, the inverse Euler equation is a common result in any dynamic Mirrlees problem, and it occurs due to uncertainties in consumption at $t = 1$ when the parent learns about their child's innate ability. If the parent can perfectly predict their child's innate ability at $t = 0$, then the standard Euler equation holds, and parents can save efficiently as if there is no government intervention. However, the inverse Euler equations tells us that the government cannot allow parents to save too much for the next period because it is more difficult to motivate parents to work optimally if they are too wealthy, and thus there are additional welfare costs from ensuring incentive compatibility in the government's problem[3]. As such, the government will optimally choose a positive distortion on personal savings for both the parent and the child.

3.0.2 Labour Wedges

The labour wedges allow us to understand the distortions on the labour supply decision in each period when the agent's productivity and ability are unobservable. Assuming that the labour market is competitive, the agent's real wage is equal to their productivity in equilibrium. Then:

[2] Proof can be found in A.0.3.

[3] More information on the inverse Euler equation can be found in Golosov, Kocherlakota, and Tsyvinski (2003).

The labour wedge at $t = 0$ for parent of type ϑ_p is defined by:

$$\tau_0^w(\vartheta_p) = 1 - \frac{h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)}{\vartheta_p u'(c_0^p(\vartheta_p))}$$

The labour wedge at $t = 1$ for parent of type ϑ_p and observes σ is defined by:

$$\tau_1^w(\sigma, \vartheta_p) = 1 - \frac{h'\left(\frac{y_1^p(\sigma,\vartheta_p)}{\vartheta_p}\right)}{\vartheta_p u'(c_1^p(\sigma, \vartheta_p))}$$

The labour wedge at $t = 2$ for child of type (σ, ϑ_k) with parent of ϑ_p is defined by:

$$\tau_2^w(\sigma, \vartheta_p, \vartheta_k) = 1 - \frac{h'\left(\frac{y_2^k(\sigma,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{\vartheta_k u'(c_2^k(\sigma, \vartheta_p, \vartheta_k))}$$

Labour supply decisions are efficient when $\tau^w = 0$, meaning that agents optimise their labour-leisure decision by equating their marginal rate of substitution between leisure and labour with the price ratio, where the price of leisure is the real wage and the price of leisure is normalised to 1. Hence when $\tau^w > 0$, agents are producing an inefficiently low supply of labour. If $\tau^w < 0$, agents are supplying a higher than efficient amount of labour.

The optimal labour wedge for the parent at $t = 0, 1$ can be separated into the following:

$$A^p(\vartheta_p) = \frac{1 - F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}$$

$$B_t^p(\vartheta_p, .) = 1 + \frac{y_t^p(\vartheta_p, .)}{\vartheta_p} \frac{h''\left(\frac{y_t^p(\vartheta_p,.)}{\vartheta_p}\right)}{h'\left(\frac{y_t^p(\vartheta_p,.)}{\vartheta_p}\right)}$$

$$C_t^p(\vartheta_p, .) = \int_{\vartheta_p}^{\overline{\vartheta}_p} \frac{u'(c_t^p(\vartheta_p, .))}{u'(c_t^p(x, .))} \left(1 - \frac{\gamma u'(c_t^p(x, .))}{\lambda}\right) \frac{f^p(x)}{1 - F^p(\vartheta_p)} dx$$

$$D_1^p(\sigma, \vartheta_p) = \int_{\vartheta_p}^{\overline{\vartheta}_p} \frac{u'(c_1^p(\sigma, \vartheta_p))}{u'(c_1^p(\sigma, x))} \left(1 - \frac{u'(c_1^p(\sigma, x))}{u'(c_0^p(x))}\right) \frac{f^p(x)}{(1 - F^p(\vartheta_p))} dx$$

Proposition 2 *The optimal marginal tax distortion for the parent at* $t = 0$ *for any* ϑ_p *and at* $t = 1$ *for any* (σ, ϑ_p) *satisfies:*

$$\frac{\tau_0^w(\vartheta_p)}{1 - \tau_0^w(\vartheta_p)} = A^p(\vartheta_p)B_0^p(\vartheta_p)C_0^p(\vartheta_p) \tag{3}$$

$$\frac{\tau_1^w(\sigma, \vartheta_p)}{1 - \tau_1^w(\sigma, \vartheta_p)} = \left[A^p(\vartheta_p)B_1^p(\sigma, \vartheta_p)\right]\left[C_1^p(\sigma, \vartheta_p) - D_1^p(\sigma, \vartheta_p)\right] \tag{4}$$

Proposition 2[4] gives us the optimal marginal tax rates for the parent whenever the government cannot observe ϑ_p and σ. The labour wedge at $t = 0$ is the standard Mirrlees-Diamond-Saez formula, otherwise known as the ABC formula. At $t = 1$, the parent receives information about their child's innate ability which creates an additional distortion D to labour supply. The labour wedge at $t = 1$ can be split into two parts: the intratemporal and intertemporal component. The intratemporal component 'ABC' at $t = 0, 1$ captures the costs and benefits of labour distortions in providing insurance against productivity shock at t. The intertemporal distortion at $t = 1$ is similar to Golosov et al. (2016), and captures the government's ability to relax the marginal tax rate to compensate for lower consumption in the previous period.

To understand these economic forces in detail, consider the labour distortion at $t = 0$. $A^p(\vartheta_p)$ represents the tail of the productivity distribution. More specifically, it measures the thickness of the right tail such that a distribution which is skewed towards high ϑ_p will result in a higher optimal marginal tax rate on a ϑ_p parent. To understand this, whenever the measure of parents above ϑ_p (i.e. $1 - F^p(\vartheta_p)$) is high relative to the measure of parents of ϑ_p type (i.e. $f^p(\vartheta_p)$), then the marginal tax rate must be high enough to disincentivise parents of $x > \vartheta_p$ from reporting as a lower productivity type. $B_0^p(\vartheta_p)$ is the Frisch elasticity of labour supply, hence if labour supply is elastic then the marginal tax rate should be lower since the cost of redistribution is too high, owing to greater losses in production. Finally, $C_0^p(\vartheta_p)$ measures the degree of inequality of the consumption allocations by the government, and how much redistribution is required by the government. To see this, the optimal tax rate is increasing in the inequality of consumption across parents as shown by the term $\frac{u'(c_0^p(\vartheta_p))}{u'(c_0^p(x))}$. Furthermore, $\left(1 - \frac{\gamma u'(c_1^p(\sigma,x))}{\lambda}\right)$ can be rewritten as $\left(1 - \frac{\mathbb{E}_{\vartheta_p}\left[\frac{1}{u'(c_0^p(\vartheta_p))}\right]}{\frac{1}{u'(c_0^p(x))}}\right)$[5] where the inverse marginal utility is the cost of a marginal

[4]Proof can be found in A.0.4.

[5]See A.0.4 for proof of this.

increase in utility. Hence, if the cost of increasing the marginal utility of an x productivity-type parent is lower than the cost of increasing the average marginal utility $\left(\frac{1}{u'(c_0^p(x))} < \mathbb{E}_{\vartheta_p}\left[\frac{1}{u'(c_0^p(\vartheta_p))}\right]\right)$, then the $C_i^p(.)$ components are positive. This is because there are positive net benefits from: (1) increasing taxes for a ϑ_p agent which then relaxes the incentive constraints for parents of $x \geq \vartheta_p$ productivity, and (2) increasing transfers towards lower productivity parents (Paluszynski and Yu, 2019). Finally, a lower γ implies that the government cares more about the child and will increase the tax rate to redistribute more.

At $t = 1$, the labour wedge can be separated into the following components:

$$\underbrace{A^p(\vartheta_p)B_1^p(\sigma,\vartheta_p)C_1^p(\sigma,\vartheta_p)}_{\text{Intratemporal Component}} - \underbrace{A^p(\vartheta_p)B_1^p(\sigma,\vartheta_p)D_1^p(\sigma,\vartheta_p)}_{\text{Intertemporal Component}}$$

At $t = 1$, the labour distortion is exposed to the same economic forces as the ones explained at $t = 0$. However, the parent now considers intertemporal trade-offs captured by $A^p(\vartheta_p)B_1^p(\sigma,\vartheta_p)D_1^p(\sigma,\vartheta_p)$. More specifically, the $D_1^p(\sigma,\vartheta_p)$ component resembles the intertemporal wedge found in Golosov et al. (2016), and this component is positive. This captures how the government uses distortions at $t = 1$ as a reward for revealing information at $t = 0$, given that the optimal intertemporal wedge at $t = 0$ mandates that the parent should consume less at $t = 1$.

Notice that the characteristics of the optimal labour wedge at $t = 1$ are independent of the child's innate ability. This is because the new information has no effect on the parent's productivity, hence the government can relax the marginal tax rate for both types of parents at $t = 1$. Furthermore, since the parent directly receives utility from the child's expected welfare, the parent will optimally choose to invest in college education for their child as long as the returns to education and future output is increasing in innate ability (See Lemma 2) without the need for additional distortions by the government.

3.0.3 Intergenerational Wedge

The optimal college savings account can be designed by characterising the intertemporal trade-offs between the parent and the child. The distortion to the intergenerational Euler equation is defined by the following intergenerational wedges. The intergenerational wedge at $t = 0$ for parent of type ϑ_p is defined by:

$$\tau_0^i(\vartheta_p) = 1 - \frac{u'(c_0^p(\vartheta_p))}{\alpha\mathbb{E}_\sigma\left[u'(c_1^k(\sigma,\vartheta_p))\right]}$$

The intergenerational wedge at $t = 1$ for parent of type ϑ_p and child of type σ is:

$$\tau_1^i(\sigma, \vartheta_p) = 1 - \frac{u'(c_1^p(\sigma, \vartheta_p))}{\alpha \mathbb{E}_\sigma \left[u'(c_2^k(\sigma, \vartheta_p, \vartheta_k)) \right]}$$

When $\tau^i = 0$, the intergenerational Euler holds and parental transfers are socially optimal. However, information asymmetry and government intervention create distortions to the efficient level of inter vivos transfers. In particular, if $\tau^i > 0$, then the parent is saving less than the efficient amount for their child. If $\tau^i < 0$, then the parent is saving more than the efficient amount for their child.

Proposition 3 *The optimal intergenerational transfers from the parent to the child at* $t = 0$ *for any* ϑ_p *and at* $t = 1$ *for any* (σ, ϑ_p) *satisfies:*

$$\frac{1}{u'(c_0^p(\vartheta_p))} = \frac{1}{\alpha} \mathbb{E}_\sigma \left[\frac{1}{u'(c_1^k(\sigma, \vartheta_p))} \right] - \frac{1-\gamma}{\alpha\lambda} \tag{7}$$

$$\frac{1}{u'(c_1^p(\sigma, \vartheta_p))} = \frac{1}{\alpha} \mathbb{E}_{\vartheta_k} \left[\frac{1}{u'(c_2^k(\sigma, \vartheta_p))} \right] - \frac{1-\gamma}{\alpha\lambda} \tag{8}$$

To understand Proposition 4[6], first consider Equation (7). The optimal intergenerational wedge at $t = 0$ implies that $\tau_0^i(\vartheta_p) > -\frac{(1-\gamma)u'(c_0^p(\vartheta_p))}{\alpha\lambda}$. When $\gamma = 1$ the intergenerational wedge at $t = 0$ satisfies the inverse intergenerational Euler equation, meaning that $\tau_0^i(\vartheta_p) > 0$. Intuitively, whenever the government does not directly care about the child's welfare, then any transfers towards the child will only indirectly affect welfare through the parent's utility via altruism. As such, the government will find it optimal to place a positive distortion on inter vivos savings, similar to the case with the parent's personal savings. However, when $\gamma < 1$, the government takes into account the benefits of the education investment on the child's future productivity and welfare, and it becomes optimal to begin subsidising inter vivos transfers. In fact, Proposition 4 provides three important intuitions on subsidised college saving accounts.

First, a higher welfare weighting on the child placed by the government relative to the parent's altruism parameter will result in a higher taxation subsidy on college savings. Alternatively, a higher altruism parameter relative to the government's weighting on child will result in a lower subsidy since the parent gains a higher utility from increasing their savings towards their child without the need for the government to help incentivise future productivity gains. Second, $\tau_0^i(\vartheta_p)$ is decreasing in $u'(c_0^p(\vartheta_p))$ for all ϑ_p, meaning that less wealthy parents in terms

[6] Proof can be found in A.0.5.

of lower consumption should be subsidised more. This is particularly interesting because this implies that the savings account should be contingent on income. Third, λ is the shadow price on the resource constraint which can be interpreted as the government's willingness to pay for an additional unit of resource[7]. Therefore, if λ is high then it is more costly to fund for the subsidy, and so the government will be less inclined to provide the subsidised college savings account.

The optimal intergeneration savings wedge at $t = 1$ follows the same ideas as $t = 0$. This result occurs because the positive personal savings distortion at $t = 1$ for the child is being offset by a transfer from the parent to the child. Interestingly, the government must offer a relative subsidy for parents with a child of low innate ability at $t = 1$. Intuitively, government must begin subsidizing transfers at $t = 0$ before the parents realise their child's innate ability. Hence, setting the optimal intergenerational wedge in the same manner for the high and low innate ability types at $t = 1$ will avoid additional distortions to incentives, given that parents would like to smooth consumption over time.

[7]Recall that $\frac{1}{\lambda} = \mathbb{E}_{\vartheta_p}\left[\frac{1}{u'(c_0^p(\vartheta_p))}\right]$ represents the cost of increasing the average marginal utility.

Chapter 4
Implementation

The optimal wedges from the previous section provides insight into policy designs on college savings and income taxes. These wedges characterise the optimal allocations that are achieved through a direct mechanism; whereby the parent and child report their innate ability and productivity types, and the government assigns an allocation on consumption, output, and investment decision as functions of the report.
This section decentralises the implementation of the optimum allocations through income taxes and separate policies for the parent's personal savings, and specialised savings for the child. The difference here is that the government is implementing policies based on the observed information of income and investment decision, rather than focusing on reports on innate ability and productivity.
First, it can be shown that the optimal consumption from the direct mechanism can be rewritten as a function of income and education. As such, the government can implement income taxes, and offer two types of education and income-contingent saving accounts to the parent; one where the parent can save in a non-specialised savings account, and one where the parent saves directly towards the child's future education[1].

Lemma 4 *At* $\mathrm{t} = 0$*, the optimal consumption* $\mathrm{c}_0^{\mathrm{p}}(\vartheta_{\mathrm{p}})$ *is a function of* $\mathrm{y}_0^{\mathrm{p}}(\vartheta_{\mathrm{p}})$ *such that* $\mathrm{c}_0^{\mathrm{p}}(\vartheta_{\mathrm{p}}) = \mathrm{c}_0^{\mathrm{p}}(\mathrm{y}_0^{\mathrm{p}}(\vartheta_{\mathrm{p}}))$*, and at* $\mathrm{t} = 1$*, for any* $\mathrm{e} \in \{0, 1\}$ *the optimal consumption* $\mathrm{c}_1^{\mathrm{i}}(\sigma, \vartheta_{\mathrm{p}})$ *are functions of education* e *and parental income* $\mathrm{y}_1^{\mathrm{p}}(\sigma, \vartheta_{\mathrm{p}})$ *such that* $\mathrm{c}_1^{\mathrm{i}}(\sigma, \vartheta_{\mathrm{p}}) = \mathrm{c}_1^{\mathrm{i}}(\mathrm{e}, \mathrm{y}_1^{\mathrm{p}}(\mathrm{e}, \vartheta_{\mathrm{p}}))$.

4.0.1 Specialised Savings Account

The simplest implementation to consider is the case where the parent has two types of savings: $\mathrm{S}_0^{\mathrm{p}}$ which can be used for the parent's own consumption at $\mathrm{t} = 1$, and $\mathrm{S}_0^{\mathrm{k}}$ which can be used for the child's consumption at $\mathrm{t} = 1$. The return on savings for these accounts are affected by income-dependent policies Γ_0^{p} and Γ_0^{k} on the parent and child's account respectively. At $\mathrm{t} = 1$, the parent can either use the savings to send the child to college or spend it on other consumption goods for the child. However, the government specifies a policy Ω_1^{k} on the use of the savings from the

[1]Note that if the conditions of Lemma 2 holds, then the parent will find it optimal for a high innate ability child to be sent to college

child's account, dependent on the investment decision made by the parent, and the wealth of the child as dictated by the parent's income. These features are inspired by the current policies on college savings, such as the 529 plan and the RESP. Both of these policies allow the parent to make contributions towards a college savings plan, which can later be withdrawn for qualified educational expenses. Finally, the parent will face some income tax $T(y_t^p)$ during their working periods at $t = 0, 1$. Given the policies described so far, the parent solves a two-stage problem. First, at $t = 0$ the parent of ϑ_p productivity solves the following:

$$\max_{c_0^p, c_1^p, c_1^k, S_1^p, S_1^k} \left\{ u(c_0^p) - h\left(\frac{y_0^p}{\vartheta_p}\right) + \beta \mathbb{E}_\sigma \left[u(c_1^p) - h\left(\frac{y_1^p}{\vartheta}\right)\right] + \alpha\beta \mathbb{E}_\sigma \left[\beta_1^e u(c_1^k) + \beta_2^e \int_{\underline{\vartheta}}^{\overline{\vartheta}} \left(u(c_2^k) - h\left(\frac{y_2^k}{\vartheta_k}\right)\right) d\vartheta_k \right] \right\}$$

subject to

$$\begin{aligned} c_0^p + S_1^p + S_1^k &= y_0^p - T(y_0^p) \\ c_1^p &= y_1^p - T(y_1^p) + (1 + \Gamma_0^p(y_0^p))RS_1^p \\ c_1^k &= (1 + \Gamma_0^k(y_0^p))RR_1^k(e)S_1^k \end{aligned}$$

Second, at $t = 1$ the parent of ϑ_p productivity have decided on whether to send the child to college. As such, the parent solves the following:

$$\max_{c_1^p, c_1^k} \left[u(c_1^p) - h\left(\frac{y_1^p}{\vartheta_p}\right) + \alpha\beta_1^e u(c_1^k) + \beta_2^e E_{\vartheta_k} U_2^k \right]$$

subject to

$$c_1^p + c_1^k(1 - \Omega_1^k(e)) = y_1^p + R(1 + \Gamma_0^p(y_0^p(\vartheta_p))S_0^p + RR_1^e(1 + \Gamma_0^k(y_0^p(\vartheta_p))S_0^k - T(y_1^p)$$

Finally, at $t = 2$, the child of $(\sigma, \vartheta_p, \vartheta_k)$ type solves:

$$\max_{c_2^k} \left\{ u(c_2^k) - h\left(\frac{y_2^k}{\vartheta_k}\right)\right\}$$

subject to

$$c_2^k = y_2^k - T(y_2^p)$$

By looking at the marginal rates of substitution in the maximisation problems, it can be shown that:

$$\Gamma_0^p(y_0^p) = u'(c_0^p)\underbrace{\left[\frac{1}{\mathbb{E}_\sigma u'(c_1^p)} - \mathbb{E}_\sigma\left(\frac{1}{u'(c_1^p)}\right)\right]}_{<0} \tag{9}$$

$$\Gamma_0^k(y_0^p) = u'(c_0^p)\underbrace{\left[\frac{1}{\alpha\mathbb{E}_\sigma u'(c_1^k)} - \mathbb{E}_\sigma\left(\frac{1}{\alpha u'(c_1^k)}\right)\right]}_{<0} + \frac{(1-\gamma)u'(c_0^p)}{\alpha\lambda} \tag{10}$$

$$\Omega_1^k(e) = \frac{(1-\gamma)u'(c_k(\sigma))}{\alpha\lambda} \tag{11}$$

Equations (9) and (10) are the wedges on the savings account that decentralise the optimum. While equation (11) is the wedge on the child's consumption at $t = 1$. The implications of this decentralisation are similar to the discussion provided in the theoretical results from the government's problem. From equations (9) and (10), the government will place a strictly positive income-dependent tax on savings on the personal savings account, and a relative subsidy on the returns to the child's savings account. Equation (11) tells us that the use of funds from these savings are dependent on the investment decision by the parent. Specifically, the subsidy on consumption at $t = 1$ is dependent on whether the child goes to college.

4.0.2 Comparison with Current Policies

The optimal wedges derived in section 3 provides insight into different types of implementation methods that can decentralise the optimal allocations. Section 4.1 provided a simple implementation to give an idea of how the government can do this. Here, I compare the simple implementation with current policies, and provide insights into other potential implementations.

The 529 plan in the U.S. is an example of a college savings account that is considered in the implementation. The 529 plan is a tax-advantaged investment vehicle provided by the government to alleviate the burdens of future rises in education expenses for a designated beneficiary. The design of the 529 plan exhibit characteristics that match the results found in the government's problem. For example, it is advantageous to begin making contributions to the 529 plan early. This relates to the theoretical results that find it optimal for the parent to set aside additional savings for the child, despite shocks to the realisation of the child's innate ability. The 529 plan does this by offering tax-deferred growth of the contributions, and tax-free withdrawals whenever it is used for qualified education expenses such as tuition fees and textbooks. Furthermore, the nature of this savings vehicle favour early savers, for example, asset allocations become more conservative as the beneficiary is closer to college age.

Another feature that is considered in my implementation is that the funds can still be used on the child, despite not going to college. Withdrawals from the 529 plan for non-qualified education expenses are allowed, but it is subject to the federal income tax alongside a 10% penalty rate on the withdrawal. The implementation here shows that it is optimal for these savings to be used by the beneficiary rather than the contributor, even after allowing for penalties for non-qualified education expenses. This suggests that the penalties for the non-qualified withdrawal by the contributor should differ from the penalties for non-qualified withdrawal by the beneficiary.

Next, we consider the RESP in Canada. The design is very similar to the 529 plan in terms of providing tax-advantages whenever contributions are made to the plan. The main difference in terms of design is the additional subsidy provided by the government. Whenever a parent or guardian makes a contribution to the RESP, the government provides an additional 20% on top of the contribution each year (up to a maximum of $500 in subsidy). This ultimately increases the returns to contribution from the RESP. However, this subsidy is also dependent on the income of the contributor. Low to middle-income families can benefit from an additional 20% on the first $500 contributed each year if their income is below a certain threshold. This feature is important in the implementation because the size of the subsidy on the college savings account in Section 4.1 should depend on the income observed by the government. A combination of the features found in the 529 plan and the RESP can implement the optimum.

Chapter 5
Discussion and Conclusion

5.0.1 Model Limitations

Although most of the assumptions in the model have been addressed throughout the paper, there are clear limitations of the results owing to these assumptions. First, a large source of financial aid for college comes in the form of student loans. This source of finance is omitted from the model because the goal of this paper is to understand how the government should distort inter vivos transfers given that we observe that parents choose to invest in their children's human capital. Although minor, it should be noted that the contributions to a 529 plan can potentially affect the amount of financial aid received by the child. However, the design of the 529 plan ensures that the contributions to this plan will not have a major effect on the child's future financial aid. For example, the plan is considered as the parent's asset rather than the beneficiary's [1].

Second, the child does not provide any labour during the education period. This means that the consumption of the child during this period is dependent only on the parent's income. As such the policies on the child's consumption before t = 2 can only be expressed in terms of parental income and education investment. This prevents us from comparing certain features relating to the child's income in current policies. For example, the 529 plan allow tax-free withdrawals, while the RESP's withdrawal for qualified education expenses are still taxed at the beneficiary's income tax rate. The lack of child income at t = 1 does not allow for this specification. However, in most cases the beneficiary's income will be low enough such that the withdrawals are essentially tax-free.

Finally, there are no disagreements in the intertemporal trade-offs between the parent and child. In this model, it is assumed that the child will go to college if the parent saves and puts money into the college savings account. This will in turn affect the child's human capital, and productivity once the child enters the workforce. However, differing views from the parent and child's optimal choice of education investment may affect policy designs. For example, is it optimal to apply the withdrawal penalties to all non-college related expenses? This is particularly interesting because the Tax Cuts and Jobs Act of 2017 expanded the 529 plan to

[1] There are also applications that help determine how much the parent is expected to contribute to the 529 plan, and how much financial aid the child receives.

allow for the savings to be used on certain non-college education expenditure such as K-12 public, private, and religious school tuition fees.

5.0.2 Calibration by Quantitative Analysis

On top of exploring how changes to the model could potentially change the optimal policy and design as mentioned in the previous sections, further exercises can be done with the results that I have derived in this paper. One major component of research that is not analysed in this paper is a calibration of the model. A calibration exercise will provide an understanding of the size of the wedges, and give a quantitative idea on what the welfare gains are from implementing the optimal policy that are consistent with the theoretical results, compared to a simple linear or any second-best optimal policy.

5.0.3 Conclusion

There are concerns over the increasing financial burden of attending college. The U.S. and Canada facilitate programs that aim at incentivising parents and family members to contribute towards a child's college savings. However, these programs differ in design. This paper attempts to address this by characterizing the optimal design of a college savings account by considering a dynamic Mirrlees model with intergenerational human capital investment. The general inverse Euler equation holds, and the labour wedges experience similar forces to most dynamic Mirrlees problems. However, the overlapping generation allow for the characterisation of redistribution in consumption between the parent and child during the education period. Finally, it is shown that the optimal design for a subsidised savings account should be contingent on parental income, and this an explicit feature in the RESP but not the 529 plan.

Appendix A
Proofs

A.0.1 Incentive Compatibility

Lemma 1:
For any σ *and* ϑ_p*,* K *is ex-post incentive compatible if and only if* ***(i)*** $y_2^k(\sigma, \vartheta_p, \vartheta_k)$ *is non-decreasing in* ϑ_k *and* ***(ii)*** $U_2^k(\sigma, \vartheta_p, \vartheta_k)$ *is absolutely continuous in* ϑ_k*, with* $\frac{\delta U_2^k(\sigma,\vartheta_p,\vartheta_k)}{\delta\vartheta_k} = \frac{y_2(\sigma,\vartheta_p,\vartheta_k)}{\vartheta_k^2} h'\left(\frac{y_2(\sigma,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)$.

Proof (i) To show that $y_2^k(\sigma, \vartheta_p, \vartheta_k)$ is non-decreasing in ϑ_k. Given (σ, ϑ_p), the incentive compatibility constraint with $\hat{\vartheta}_k > \vartheta_k$ can be written as follows:

$$(1) \quad u(c_2^k(\sigma, \vartheta_p, \vartheta_k)) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\vartheta_k}\right) \geq u(c_2^k(\sigma, \vartheta_p, \hat{\vartheta}_k)) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \hat{\vartheta}_k)}{\vartheta_k}\right)$$

$$(2) \quad u(c_2^k(\sigma, \vartheta_p, \hat{\vartheta}_k)) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \hat{\vartheta}_k)}{\hat{\vartheta}_k}\right) \geq u(c_2^k(\sigma, \vartheta_p, \vartheta_k)) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\hat{\vartheta}_k}\right)$$

Adding (1) and (2) and simplifying bothing sides will obtain:

$$h\left(\frac{y_2^k(\sigma, \vartheta_p, \hat{\vartheta}_k)}{\vartheta_k}\right) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \hat{\vartheta}_k)}{\hat{\vartheta}_k}\right) \geq h\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\vartheta_k}\right) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\hat{\vartheta}_k}\right)$$

Now, define:

$$g(y_2^k) = h\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta)}{\vartheta_k}\right) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta)}{\hat{\vartheta}_k}\right)$$

Taking the first derivative:

$$g'(y_2^k) = \frac{1}{\vartheta_k} h'\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta)}{\vartheta_k}\right) - \frac{1}{\hat{\vartheta}_k} h'\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta)}{\hat{\vartheta}_k}\right)$$

but since $\hat{\vartheta}_k > \vartheta_k$, it must be that $g'(y_2^k) > 0$ so that $y_2^k(\sigma, \vartheta_p, \hat{\vartheta}_k) > y_2^k(\sigma, \vartheta_p, \vartheta_k)$ and y_2^k is increasing in ϑ_k

■

Proof (ii) Given (σ, ϑ_p) the ex-post ICC can be expressed as:

$$\max_{\hat{\vartheta}_k} \left\{ u(c_2^k(\sigma, \vartheta_p, \hat{\vartheta}_k)) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \hat{\vartheta}_k))}{\vartheta_k}\right) \right\}$$

Where ϑ_k is the maximising $\hat{\vartheta}_k$. Hence the FOC is:

$$\frac{\delta u(c_2^k(\sigma, \vartheta_p, \vartheta_k))}{\delta c_2^k(\sigma, \vartheta_p, \vartheta_k)} \frac{\delta c_2^k(\sigma, \vartheta_p, \vartheta_k)}{\delta \vartheta_k} - \left[\frac{1}{\vartheta_k} \frac{\delta y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\delta \vartheta_k} h'\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\vartheta_k}\right)\right] = 0$$

The value function at $t = 2$ is:

$$U_2^k(\sigma, \vartheta_p, \vartheta_k) = u(c_2^k(\sigma, \vartheta_p, \vartheta_k)) - h\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\vartheta_k}\right)$$

Taking the first derivative:

$$\frac{dU_2^k(\sigma, \vartheta_p, \vartheta_k)}{d\vartheta_k} = \frac{\delta u(c_2^k(\sigma, \vartheta_p, \vartheta_k)}{\delta c_2^k(\sigma, \vartheta_p, \vartheta_k)} \frac{\delta c_2^k(\sigma, \vartheta_p, \vartheta_k)}{\delta \vartheta_k} - \left[\left(\frac{1}{\vartheta_k} \frac{\delta y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\delta \vartheta_k} - \frac{y_2(\sigma, \vartheta_p, \vartheta_k)}{\vartheta_k^2}\right) .h'\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\vartheta_k}\right)\right]$$

Given the FOC and using the Envelope Theorem;

$$\frac{dU_2^k(\sigma, \vartheta_p, \vartheta_k)}{d\vartheta_k} = \frac{y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\vartheta_k^2} h'\left(\frac{y_2^k(\sigma, \vartheta_p, \vartheta_k)}{\vartheta_k}\right)$$

Therefore, the revenue equivalence is:

$$U_2^k(\sigma, \vartheta_p, \vartheta_k) = U_2^k(\sigma, \vartheta_p, \underline{\vartheta}_k) + \int_{\underline{\vartheta}}^{\vartheta} \frac{y_2^k(\sigma, \vartheta_p, k)}{k^2} h'\left(\frac{y_2^k(\sigma, \vartheta_p, k)}{k}\right) dk$$

■

Lemma 2:
Suppose K *is ex-post incentive compatible. Then if* ***(i)*** *The returns to education for high innate ability is greater than for low innate ability agents such that:*

$$\beta_2^{e_1} |F(\vartheta_k|\varkappa(H, e = 1) - F(\vartheta_k|\varkappa(L, e = 1))| - \beta_2^{e_0} |F(\vartheta_k|\varkappa(H, e = 0) - F(\vartheta_k|\varkappa(L, e = 0))| \geq 0$$

and ***(ii)*** *Output is increasing in innate ability for all* ϑ_p, ϑ_k*:* $y_2(H, \vartheta_p, \vartheta_k) \geq y_2(L, \vartheta_p, \vartheta_k)$*, then* $U_1^p(H, e = 1; \vartheta_p) = U_1^p(H, e = 0; \vartheta_p)$ *at the optimum.*

Proof The downward ex-ante ICC 2 is given as:

$$u(c_1^p(H,\vartheta_p)) - h\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right) + \alpha U(c_1^k, c_2^k, y_2^k : H, e=1)$$
$$\geq u(c_1^p(L,\vartheta_p)) - h\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right) + \alpha U(c_1^k, c_2^k, y_2^k : H, e=0)$$

where:

$$U_1^k\left(c_1^k, c_2^k, y_2^k; \sigma, e\right) = \beta_1^e u\left(c_1^k\right) + \beta_2^e \int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[u\left(c_2^k\right) - h\left(\frac{y_2^k}{\vartheta_k}\right)\right] f^k\left(\vartheta_k | \varkappa\left(\sigma, e\right)\right) d\vartheta_k.$$

Let:

$$(1) \quad \Phi(\sigma,\vartheta_p) = u(c_1^p(\sigma,\vartheta_p)) - h\left(\frac{y_1^p(\sigma,\vartheta_p)}{\vartheta_p}\right) + \alpha\beta_1^e u(c_1^k(\sigma,\vartheta_p))$$

$$(2) \quad U_2^k(\sigma,\vartheta_p,\vartheta_k) = u(c_2^k(\sigma,\vartheta_p,\vartheta_k)) - h\left(\frac{y_2^k(\sigma,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)$$

Therefore the high innate ability's ex-ante incentive constraint can be simplified to:

$$\Phi(H,\vartheta_p) + \alpha\beta_2^1 \int_{\underline{\vartheta}}^{\overline{\vartheta}} U_2^k(H,\vartheta_p,\vartheta_k) f^k\left(\vartheta_k | \varkappa\left(H, e=1\right)\right) d\vartheta_k$$
$$\geq \Phi(L,\vartheta_p) + \alpha\beta_2^0 \int_{\underline{\vartheta}}^{\overline{\vartheta}} U_2^k(L,\vartheta_p,\vartheta_k) f^k\left(\vartheta_k | \varkappa\left(H, e=0\right)\right) d\vartheta_k.$$

Since the Low-Ability ICC is symmetric, adding the High and Low incentive constraints gives:

$$\beta_2^0 \int_{\underline{\vartheta}}^{\overline{\vartheta}} U_2^k(L,\vartheta_p,\vartheta_k) f^k\left(\vartheta_k | \varkappa\left(L, e=0\right)\right) d\vartheta_k - \beta_2^0 \int_{\underline{\vartheta}}^{\overline{\vartheta}} U_2^k(L,\vartheta_p,\vartheta_k) f^k\left(\vartheta_k | \varkappa\left(H, e=0\right)\right) d\vartheta_k$$
$$+\beta_2^1 \int_{\underline{\vartheta}}^{\overline{\vartheta}} U_2^k(H,\vartheta_p,\vartheta_k) f^k\left(\vartheta_k | \varkappa\left(H, e=1\right)\right) d\vartheta_k - \beta_2^1 \int_{\underline{\vartheta}}^{\overline{\vartheta}} U_2^k(H,\vartheta_p,\vartheta_k) f^k\left(\vartheta_k | \varkappa\left(L, e=1\right)\right) d\vartheta_k \geq 0$$

Adding and subtracting: $\beta_2^0 \int_{\underline{\vartheta}}^{\overline{\vartheta}} U_2^k(H,\vartheta_p,\vartheta_k) f^k\left(\vartheta_k | \varkappa\left(L, e=0\right)\right) d\vartheta_k$ and $\beta_2^0 \int_{\underline{\vartheta}}^{\overline{\vartheta}} U_2^k(H,\vartheta_p,\vartheta_k) f^k\left(\vartheta_k | \varkappa\left(H, e=0\right)\right) d\vartheta_k$ to the expression above and factorizing yields:

$$\beta_2^0 \int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[U_2^k(H,\vartheta_p,\vartheta_k) - U_2^k(L,\vartheta_p,\vartheta_k)\right]\left[f(\vartheta_k|H,e=0) - f(\vartheta_k|L,e=0)\right]d\vartheta_k$$

$$+\beta_2^1 \int_{\underline{\vartheta}}^{\overline{\vartheta}} U_2^k(H,\vartheta_p,\vartheta_k)\left[f(\vartheta_k|H,e=1) - \frac{\beta_2^0}{\beta_2^1}f(\vartheta_k|H,e=0) - f(\vartheta_k|L,e=1) + \frac{\beta_2^0}{\beta_2^1}f(\vartheta_k|L,e=0)\right]d\vartheta_k \geq 0$$

Integration by parts and simplifying:

$$-\int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[\frac{\delta\left[U_2^k(H,\vartheta_p,\vartheta_k) - U_2^k(L,\vartheta_p,\vartheta_k)\right]}{\delta\vartheta_k}\right]\left[F(\vartheta_k|H,e=0) - F(\vartheta_k|L,e=0)\right]d\vartheta_k$$

$$-\int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[F(\vartheta_k|H,e=1) - \frac{\beta_2^0}{\beta_2^1}F(\vartheta_k|H,e=0) - F(\vartheta_k|L,e=1) + \frac{\beta_2^0}{\beta_2^1}F(\vartheta_k|L,e=0)\right]\left[U_2^{k\prime}(H,\vartheta_p,\vartheta_k)\right]d\vartheta_k \geq 0$$

or,

$$\int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[\frac{\delta\left[U_2^k(H,\vartheta_p,\vartheta_k) - U_2^k(L,\vartheta_p,\vartheta_k)\right]}{\delta\vartheta_k}\right]\left[F(\vartheta_k|H,e=0) - F(\vartheta_k|L,e=0)\right]d\vartheta_k$$

$$+\int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[F(\vartheta_k|H,e=1) - \frac{\beta_2^0}{\beta_2^1}F(\vartheta_k|H,e=0) - F(\vartheta_k|L,e=1) + \frac{\beta_2^0}{\beta_2^1}F(\vartheta_k|L,e=0)\right]\left[U_2^{k\prime}(H,\vartheta_p,\vartheta_k)\right]d\vartheta_k \leq 0$$

Using the expression for $\frac{\delta U_2^k(\sigma,\vartheta_p,\vartheta_k)}{\delta\vartheta_k}$ (Lemma 1):

$$\int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[\frac{y_2(H,\vartheta_p,\vartheta_k)}{\vartheta_k^2}h'\left(\frac{y_2(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right) - \frac{y_2(L,\vartheta_p,\vartheta_k)}{\vartheta_k^2}h'\left(\frac{y_2(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)\right]$$
$$\left[F(\vartheta_k|H,e=0) - F(\vartheta_k|L,e=0)\right]d\vartheta_k$$
$$+\int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[F(\vartheta_k|H,e=1) - \frac{\beta_2^0}{\beta_2^1}F(\vartheta_k|H,e=0) - F(\vartheta_k|L,e=1) + \frac{\beta_2^0}{\beta_2^1}F(\vartheta_k|L,e=0)\right]$$
$$\left[\frac{y_2(H,\vartheta_p,\vartheta_k)}{\vartheta_k^2}h'\left(\frac{y_2(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)\right]d\vartheta_k \leq 0$$

Given (i.) $y_2^k(H,\vartheta_p,\vartheta_k) > y_2^k(L,\vartheta_p,\vartheta_k)$, (ii.) $F(\vartheta_k|H,e=0) - F(\vartheta_k|L,e=0) < 0$, and (iii.) h(.) is convex:

$$\text{(a)} \quad \int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[\frac{y_2(H,\vartheta_p,\vartheta_k)}{\vartheta_k^2} h'\left(\frac{y_2(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right) - \frac{y_2(L,\vartheta_p,\vartheta_k)}{\vartheta_k^2} h'\left(\frac{y_2(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right) \right] \left[F(\vartheta_k|H,e=0) - F(\vartheta_k|L,e=0) \right] d\vartheta \leq 0$$

As long as (i.), (ii.), and (iii.) holds, then (a) is satisfied.

$$\text{(b)} \quad \int_{\underline{\vartheta}}^{\overline{\vartheta}} \left[F(\vartheta_k|H,e=1) - \frac{\beta_2^0}{\beta_2^1} F(\vartheta_k|H,e=0) - F(\vartheta_k|L,e=1) + \frac{\beta_2^0}{\beta_2^1} F(\vartheta_k|L,e=0) \right] \left[\frac{y_2(H,\vartheta_p,\vartheta_k)}{\vartheta_k^2} h'\left(\frac{y_2(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right) \right] d\vartheta \leq 0$$

The expression for (b) is satisfied if:

$$\beta_2^1 \left| F(\vartheta_k|\varkappa(H,e=1) - F(\vartheta_k|\varkappa(L,e=1)) \right| - \beta_2^0 \left| F(\vartheta_k|\varkappa(H,e=0) - F(\vartheta_k|\varkappa(L,e=0)) \right| \geq 0$$

Therefore if:

1. $y_2^k(H,\vartheta_p,\vartheta_k) > y_2^k(L,\vartheta_p,\vartheta_k) \qquad \forall \vartheta_p, \vartheta_k$

2. $\beta_2^1 \left| F(\vartheta_k|\varkappa(H,e=1) - F(\vartheta_k|\varkappa(L,e=1)) \right| - \beta_2^0 \left| F(\vartheta_k|\varkappa(H,e=0) - F(\vartheta_k|\varkappa(L,e=0)) \right| \geq 0 \qquad \forall \vartheta_k$

Then ex-ante ICC 2 (at $t = 1$) for $\sigma = H$ binds. ■

Lemma 3: *Suppose that K is ex-post incentive compatible and that* $U_1^p(H, e = 1; \vartheta_p) = U_1^p(L, e = 0; \vartheta_p)$ *at the optimum. Then (P,K) is ex-ante incentive compatible if and only if* ***(i)*** $y_0^p(\vartheta_p)$ *is non-decreasing in* ϑ_p *and* ***(ii)*** $U_0^p(\vartheta_p)$ *is absolutely continuous in* ϑ_p *with:*

$$\frac{\delta U_1^p(\vartheta_p)}{\delta \vartheta_p} = \frac{y_0^p(\vartheta_p)}{\vartheta_p^2} h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right) + \beta_1^p g(\vartheta_p) \left[\frac{y_1^p(H,\vartheta_p)}{\vartheta_p^2} h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right) \right]$$
$$+ \beta_1^p (1 - g(\vartheta_p)) \left[\frac{y_1^p(L,\vartheta_p)}{\vartheta_p^2} h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right) \right] + \alpha \beta_1^p g'(\vartheta_p) \left[U_1^k(H,e=0) - U_1^k(L,e=0) \right]$$

Proof Let the true productivity value of the parent be ϑ_p. Then, the incentive constraint at $t = 1$ satisfies:

$$\vartheta_p = \arg\max_{\hat{\vartheta}_p} \Bigg\{ u(c_0(\hat{\vartheta}_p)) - h\left(\frac{y_0^p(\hat{\vartheta}_p)}{\vartheta_p}\right) + \beta_1^p g(\vartheta_p)\left[u(c_1^p(\hat{\vartheta}_p, H)) - h\left(\frac{y_0^p(\hat{\vartheta}_p, H)}{\vartheta_p}\right) + \alpha U_1^k(\hat{\vartheta}_p, H, 1; \vartheta_p)\right]$$
$$+ \beta_1^p (1 - g(\vartheta_p))\left[u(c_1^p(\hat{\vartheta}_p, L)) - h\left(\frac{y_0^p(\hat{\vartheta}_p, L)}{\vartheta_p}\right) + \alpha U_1^k(\hat{\vartheta}_p, L, 0; \vartheta_p)\right]\Bigg\}$$

And the first order condition is:

$$u'(c_0(\vartheta_p))\frac{\partial c_0(\vartheta_p)}{\partial \vartheta_p} - \frac{h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)}{\vartheta_p}\frac{\partial y_0^p(\vartheta_p)}{\partial \vartheta_p} + \beta_1^p g(\vartheta_p)\left[u'(c_1^p(\vartheta_p, H))\frac{\partial c_1^p(\vartheta_p, H)}{\partial \vartheta_p} - \frac{h'\left(\frac{y_1^p(\vartheta_p, H)}{\vartheta_p}\right)}{\vartheta_p}\frac{\partial y_1^p(\vartheta_p, H)}{\partial \vartheta_p}\right.$$
$$\left. + \alpha\frac{\partial U_1^k(\vartheta_p, H, 1)}{\partial \vartheta_p}\right] + \beta_1^p(1 - g(\vartheta_p))\left[u'(c_1^p(\vartheta_p, L))\frac{\partial c_1^p(\vartheta_p, L)}{\partial \vartheta_p} - \frac{h'\left(\frac{y_1^p(\vartheta_p, L)}{\vartheta_p}\right)}{\vartheta_p}\frac{\partial y_1^p(\vartheta_p, L)}{\partial \vartheta_p} + \alpha\frac{\partial U_1^K(\vartheta_p, L, 0)}{\partial \vartheta_p}\right] = 0$$

The value function at $t = 0$ is:

$$U_0^p(\vartheta_p) = u(c_0(\vartheta_p)) - h\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right) + \beta_1^p g(\vartheta_p)\left[u(c_1^p(\vartheta_p, H)) - h\left(\frac{y_0^p(\vartheta_p, H)}{\vartheta_p}\right) + \alpha U_1^k(\vartheta_p, H, 1)\right]$$
$$+ \beta_1^p(1 - g(\vartheta_p))\left[u(c_1^p(\vartheta_p, L)) - h\left(\frac{y_0^p(\vartheta_p, L)}{\vartheta_p}\right) + \alpha U_1^k(\vartheta_p, L, 0)\right]$$

Taking the first order derivative wrt ϑ_p:

$$\frac{\partial U_0^p(\vartheta_p)}{\partial \vartheta_p} = u'(c_0(\vartheta_p))\frac{\partial c_0(\vartheta_p)}{\partial \vartheta_p} - \frac{h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)}{\vartheta_p}\frac{\partial y_0^p(\vartheta_p)}{\partial \vartheta_p} + \frac{y_0^p(\vartheta_p)}{\vartheta_p^2}h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)$$
$$+\beta_1^p g(\vartheta_p)\left[u'(c_1^p(\vartheta_p, H))\frac{\partial c_1^p(\vartheta_p, H)}{\partial \vartheta_p} - \frac{h'\left(\frac{y_1^p(\vartheta_p, H)}{\vartheta_p}\right)}{\vartheta_p}\frac{\partial y_1^p(\vartheta_p, H)}{\partial \vartheta_p} + \frac{y_1^p(\vartheta_p, H)}{\vartheta_p^2}h'\left(\frac{y_1^p(\vartheta_p, H)}{\vartheta_p}\right) + \alpha\frac{\partial U_1^K(\vartheta_p, H, 1)}{\partial \vartheta_p}\right]$$
$$+\beta_1^p(1 - g(\vartheta_p))\left[u'(c_1^p(\vartheta_p, L))\frac{\partial c_1^p(\vartheta_p, L)}{\partial \vartheta_p} - \frac{h'\left(\frac{y_1^p(\vartheta_p, L)}{\vartheta_p}\right)}{\vartheta_p}\frac{\partial y_1^p(\vartheta_p, L)}{\partial \vartheta_p} + \frac{y_1^p(\vartheta_p, L)}{\vartheta_p^2}h'\left(\frac{y_1^p(\vartheta_p, L)}{\vartheta_p}\right) + \alpha\frac{\partial U_1^K(\vartheta_p, L, 0)}{\partial \vartheta_p}\right]$$
$$+ \beta_1^p g'(\vartheta_p)\left[u(c_1^p(\vartheta_p, H)) - h\left(\frac{y_0^p(\vartheta_p, H)}{\vartheta_p}\right) + \alpha U_1^k(\vartheta_p, H, 1)\right] -$$
$$\beta_1^p g'(\vartheta_p)\left[u(c_1^p(\vartheta_p, L)) - h\left(\frac{y_0^p(\vartheta_p, L)}{\vartheta_p}\right) + \alpha U_1^k(\vartheta_p, L, 0)\right]$$

By Envelope Theorem, this is simplified to:

$$\frac{\partial U_0^p(\vartheta_p)}{\partial \vartheta_p} = \frac{y_0^p(\vartheta_p)}{\vartheta_p^2} h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right) + \beta_1^p g(\vartheta_p)\left[\frac{y_1^p(\vartheta_p, H)}{\vartheta_p^2} h'\left(\frac{y_1^p(\vartheta_p, H)}{\vartheta_p}\right)\right]$$

$$+\beta_1^p(1-g(\vartheta_p))\left[\frac{y_1^p(\vartheta_p, L)}{\vartheta_p^2} h'\left(\frac{y_1^p(\vartheta_p, L)}{\vartheta_p}\right)\right] + \beta_1^p g'(\vartheta_p)\left[u(c_1^p(\vartheta_p, H)) - h\left(\frac{y_0^p(\vartheta_p, H)}{\vartheta_p}\right) + \alpha U_1^k(\vartheta_p, H, 1)\right]$$

$$-\beta_1^p g'(\vartheta_p)\left[u(c_1^p(\vartheta_p, L)) - h\left(\frac{y_0^p(\vartheta_p, L)}{\vartheta_p}\right) + \alpha U_1^k(\vartheta_p, L, 0)\right]$$

Adding and subtracting $\alpha U_1^K(\vartheta_p, H, 0)$ and by Lemma 2, this simplifies to:

$$\frac{\partial U_0^p(\vartheta_p)}{\partial \vartheta_p} = \frac{y_0^p(\vartheta_p)}{\vartheta_p^2} h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right) + \beta_1^p g(\vartheta_p)\left[\frac{y_1^p(\vartheta_p, H)}{\vartheta_p^2} h'\left(\frac{y_1^p(\vartheta_p, H)}{\vartheta_p}\right)\right]$$

$$+\beta_1^p(1-g(\vartheta_p))\left[\frac{y_1^p(\vartheta_p, L)}{\vartheta_p^2} h'\left(\frac{y_1^p(\vartheta_p, L)}{\vartheta_p}\right)\right] + \beta_1^p g'(\vartheta_p)\Big[\alpha U_1^K(\vartheta_p, H, 0) - \alpha U_1^K(\vartheta_p, L, 0)\Big]$$

■

A.0.2 Solving the Planning Problem

To solve for the optimal conditions for the social planner, I adopt the Hamiltonian technique by solving the constrained welfare criterian subject to the 7 constraints.

Contrained Welfare Criterion:

$$\max\left\{\gamma\int_{\underline{\vartheta}}^{\overline{\vartheta}}U_0^p(\vartheta_p)dF(\vartheta_p)+(1-\gamma)\beta_1^p\left[\int_{\underline{\vartheta}}^{\overline{\vartheta}}g(\vartheta_p)U_1^k(H,1)dF(\vartheta_p)+\int_{\underline{\vartheta}}^{\overline{\vartheta}}(1-g(\vartheta_p))U_1^k(L,0)dF(\vartheta_p)\right]\right\}$$

(1) Resource constaint (λ):

$$\int_{\underline{\vartheta}_p}^{\overline{\vartheta}_p}\left[y_0^p(\vartheta_p)-c_0^p(\vartheta_p)+\frac{g(\vartheta_p)}{R_1^p}\left[y_1^p(H,\vartheta_p)-c_1^p(H,\vartheta_p)-\left(\frac{c_1^k(H,\vartheta_p)+e_H}{R_1^k(1)}\right)\right.\right.$$
$$\left.+\frac{1}{R_2^k(1)}\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k}\left(y_2^k(H,\vartheta_p,\vartheta_k)-c_2^k(H,\vartheta_p,\vartheta_k)\right)dF^k(\vartheta_k|H,1))\right]+\frac{1-g(\vartheta_p)}{R_1^p}\left[y_1^p(L,\vartheta_p)-c_1^p(L,\vartheta_p)-\right.$$
$$\left.\left.\left(\frac{c_1^k(L,\vartheta_p)+e_L}{R_1^k(0)}\right)+\frac{1}{R_2^k(0)}\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k}\left(y_2^k(L,\vartheta_p,\vartheta_k)-c_2^k(L,\vartheta_p,\vartheta_k)\right)dF^k(\vartheta_k|L,0))\right]\right]dF(\vartheta_p)$$

(2) Discounted Life-time utility of parent ($\mu(\vartheta_p)$)

$$U_0^p(\vartheta_p)=u(c_0^p(\vartheta_p))-h\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)+\beta g(\vartheta_p)\left[u(c_1^p(H,\vartheta_p))-h\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)+\alpha U_1^k(H,1)\right]$$
$$+\beta(1-g(\vartheta_p))\left[u(c_1^p(L,\vartheta_p))-h\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)+\alpha U_1^k(L,0)\right]$$

(3) Incentive Constraint $t=0$ parent's productivity ($\pi(\vartheta_p)$)

$$\frac{\partial U_0^p(\vartheta_p)}{\partial\vartheta_p}=\frac{y_0^p(\vartheta_p)}{\vartheta_p^2}h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)+\beta g(\vartheta_p)\left[\frac{y_1^p(H,\vartheta_p)}{\vartheta_p^2}h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)\right]$$
$$+\beta(1-g(\vartheta_p))\left[\frac{y_1^p(L,\vartheta_p)}{\vartheta_p^2}h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)\right]+\beta g'(\vartheta_p)\left[u(c_1^p(H,\vartheta_p))-h\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)+\alpha U_1^k(H,1)\right]$$
$$+\beta g'(\vartheta_p)\left[u(c_1^p(L,\vartheta_p))-h\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)+\alpha U_1^k(L,0)\right]$$

(4) Incentive Constraint t = 1 Child's type ($\rho(\vartheta_p)$)

$$u(c_1^p(H,\vartheta_p)) - h\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right) + \alpha U_1^k(H,1) = u(c_1^P(L,\vartheta_p)) + h\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)$$
$$-\alpha\left[\beta_1^0 u(c_1^k(L,\vartheta_p)) + \beta_2^0\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k}\left[u(c_2^k(L,\vartheta_p,\vartheta_k)) - h\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)\right] dF^k(\vartheta_k|H,0)\right]$$

(5) Discounted Life-time utility of child ($\omega(\sigma,\vartheta_p)$)

$$U_1^k(\sigma,e) = \beta_1^e u(c_1^k(\sigma,\vartheta_p)) + \beta_2^e\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} U_2^k(\sigma,\vartheta_p,\vartheta_k)dF^k(\vartheta_k|\sigma,e)$$

(6) Child's utility at t = 2 ($\varphi(\sigma,\vartheta_p,\vartheta_k)$)

$$U_2^k(\sigma,\vartheta_p,\vartheta_k) = u(c_2^k(\sigma,\vartheta_p,\vartheta_k)) - h\left(\frac{y_2^k(\sigma,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)$$

(7) Incentive Constraint t = 2 Child's productivity ($\psi(\sigma,\vartheta_p,\vartheta_k)$)

$$\frac{\partial U_2^k(\sigma,\vartheta_p,\vartheta_k)}{\partial\vartheta_k} = \frac{y_2^k(\sigma,\vartheta_p,\vartheta_k)}{\vartheta_k^2}h'\left(\frac{y_2^k(\sigma,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)$$

First Order Conditions at t = 0 and t = 1:

$$\frac{\partial \mathcal{L}}{\partial U_0^p} : \frac{\partial \pi(\vartheta_p)}{\partial \vartheta_p} = \gamma f^p(\vartheta_p) - \mu(\vartheta_p) \tag{A.1}$$

$$\frac{\partial \mathcal{L}}{\partial c_0^p} : \lambda f^p(\vartheta_p) = \mu(\vartheta_p) u'(c_0^p(\vartheta_p)) \tag{A.2}$$

$$\frac{\partial \mathcal{L}}{\partial y_0^p} : \lambda f^p(\vartheta_p) = \mu(\vartheta_p) \left[\frac{1}{\vartheta_p} h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)\right] + \pi(\vartheta_p) \left[\frac{1}{\vartheta_p^2} h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right) + \frac{y_0^p(\vartheta_p)}{\vartheta_p^3} h''\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)\right] \tag{A.3}$$

$$\frac{\partial \mathcal{L}}{\partial U_1^k(H)} : \omega(H, \vartheta_p) = (1-\gamma)\beta g(\vartheta_p) f^p(\vartheta_p) + \mu(\vartheta_p)\alpha\beta g(\vartheta_p) - \pi(\vartheta_p)\alpha\beta g'(\vartheta_p) + \rho(\vartheta_p)\alpha \tag{A.4}$$

$$\frac{\partial \mathcal{L}}{\partial U_1^k(L)} : \omega(L, \vartheta_p) = (1-\gamma)\beta(1-g(\vartheta_p)) f^p(\vartheta_p) + \mu(\vartheta_p)\alpha\beta(1-g(\vartheta_p)) + \pi(\vartheta_p)\alpha\beta g'(\vartheta_p) \tag{A.5}$$

$$\frac{\partial \mathcal{L}}{\partial c_1^p(H)} : \frac{\lambda g(\vartheta_p) f^p(\vartheta_p)}{R_1^p} = \Big[\mu(\vartheta_p)\beta g(\vartheta_p) - \pi(\vartheta_p)\beta g'(\vartheta_p) + \rho(\vartheta_p)\Big] u'(c_1^p(H, \vartheta_p)) \tag{A.6}$$

$$\frac{\partial \mathcal{L}}{\partial c_1^p(L)} : \frac{\lambda (1-g(\vartheta_p)) f^p(\vartheta_p)}{R_1^p} = \Big[\mu(\vartheta_p)\beta(1-g(\vartheta_p)) + \pi(\vartheta_p)\beta g'(\vartheta_p) - \rho(\vartheta_p)\Big] u'(c_1^p(L, \vartheta_p)) \tag{A.7}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial y_1^p(H)} : \frac{\lambda g(\vartheta_p) f^p(\vartheta_p)}{R_1^p} &= \Big[\mu(\vartheta_p)\beta g(\vartheta_p) - \pi(\vartheta_p)\beta g'(\vartheta_p) + \rho(\vartheta_p)\Big] \frac{1}{\vartheta_p} h'\left(\frac{y_1^p(H, \vartheta_p)}{\vartheta_p}\right) \\ &+ \pi(\vartheta_p)\left[\beta g(\vartheta_p)\left(\frac{1}{\vartheta_p^2} h'\left(\frac{y_1^p(H, \vartheta_p)}{\vartheta_p}\right) + \frac{y_1^p(H, \vartheta_p)}{\vartheta_p^3} h''\left(\frac{y_1^p(H, \vartheta_p)}{\vartheta_p}\right)\right)\right] \end{aligned} \tag{A.8}$$

$$\frac{\partial\mathcal{L}}{\partial y_1^p(L)} : \frac{\lambda(1-g(\vartheta_p))f^p(\vartheta_p)}{R_1^p} = \Big[\mu(\vartheta_p)\beta(1-g(\vartheta_p))+\pi(\vartheta_p)\beta g'(\vartheta_p)-\rho(\vartheta_p)\Big]\frac{1}{\vartheta_p}h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)$$

$$+\pi(\vartheta_p)\left[\beta(1-g(\vartheta_p))\left(\frac{1}{\vartheta_p^2}h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)+\frac{y_1^p(L,\vartheta_p)}{\vartheta_p^3}h''\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)\right)\right] \quad (A.9)$$

$$\frac{\partial\mathcal{L}}{\partial c_1^k(H)} : \frac{\lambda g(\vartheta_p)f^p(\vartheta_p)}{R_1^pR_1^k(1)} = \omega(H,\vartheta_p)\beta_1^1u'(c_1^k(H,\vartheta_p)) \quad (A.10)$$

$$\frac{\partial\mathcal{L}}{\partial c_1^k(L)} : \frac{\lambda(1-g(\vartheta_p))f^p(\vartheta_p)}{R_1^pR_1^k(0)} = \Big[\omega(L,\vartheta_p)\beta_1^0-\rho(\vartheta_p)\alpha\beta_1^0\Big]u'(c_1^k(L,\vartheta_p)) \quad (A.11)$$

First Order Conditions at t = 2:

$$\frac{\partial\mathcal{L}}{\partial c_2^k(H)} : \frac{\lambda g(\vartheta_p)f^k(\vartheta_k|H,1)f^p(\vartheta_p)}{R_1^pR_2^k(1)} = \varphi(H,\vartheta_p,\vartheta_k)u'(c_2^k(H,\vartheta_p,\vartheta_k)) \quad (A.12)$$

$$\frac{\partial\mathcal{L}}{\partial c_2^k(L)} : \frac{\lambda(1-g(\vartheta_p))f^k(\vartheta_k|L,0)f^p(\vartheta_p)}{R_1^pR_2^k(0)} = \Big[\varphi(L,\vartheta_p,\vartheta_k)-\rho(\vartheta_p)\alpha\beta_2^0f^k(\vartheta_k|H,0)\Big]u'(c_2^k(L,\vartheta_p,\vartheta_k)) \quad (A.13)$$

$$\frac{\partial\mathcal{L}}{\partial y_2^K(H)} : \frac{\lambda g(\vartheta_p)f^k(\vartheta_k|H,1)f^p(\vartheta_p)}{R_1^pR_2^k(1)} = \varphi(H,\vartheta_p,\vartheta_k)\frac{1}{\vartheta_k}h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)$$

$$+\psi(H,\vartheta_p,\vartheta_k)\left[\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k^3}h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)+\frac{1}{\vartheta_k^2}h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)\right] \quad (A.14)$$

$$\frac{\partial\mathcal{L}}{\partial y_2^K(L)} : \frac{\lambda(1-g(\vartheta_p))f^k(\vartheta_k|L,0)f^p(\vartheta_p)}{R_1^pR_2^k(0)} = \Big[\varphi(L,\vartheta_p,\vartheta_k)-\rho(\vartheta_p)\alpha\beta_2^0f^k(\vartheta_k|H,0)\Big]$$

$$\frac{1}{\vartheta_k}h'\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)+\psi(L,\vartheta_p,\vartheta_k)\left[\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k^3}h''\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)+\frac{1}{\vartheta_k^2}h'\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)\right] \quad (A.15)$$

$$\frac{\partial\mathcal{L}}{\partial U_2^k(H)} : \frac{\partial\psi(H,\vartheta_p,\vartheta_k)}{\partial\vartheta_k} = \omega(H,\vartheta_p)\beta_2^1f^k(\vartheta_k|H,1)-\varphi(H,\vartheta_p,\vartheta_k) \quad (A.16)$$

$$\frac{\partial \mathcal{L}}{\partial \mathrm{U}_2^{\mathrm{k}}(\mathrm{L})} : \frac{\partial \psi(\mathrm{L}, \vartheta_{\mathrm{p}}, \vartheta_{\mathrm{k}})}{\partial \vartheta_{\mathrm{k}}} = \omega(\mathrm{L}, \vartheta_{\mathrm{p}}) \beta_2^0 \mathrm{f}^{\mathrm{k}}(\vartheta_{\mathrm{k}} | \mathrm{L}, 0) - \varphi(\mathrm{L}, \vartheta_{\mathrm{p}}, \vartheta_{\mathrm{k}}) \tag{A.17}$$

A.0.3 Proof of Proposition 1

Intertemporal Wedge: Parent $t = 0$ and $t = 1$:

Since $\beta_1^p R_1^p = 1$, FOC (6) and (7) can be rewritten:

$$\lambda f^p(\vartheta_p) = \left[\mu(\vartheta_p) - \pi(\vartheta_p)\frac{g'(\vartheta_p)}{g(\vartheta_p)} + \rho(\vartheta_p)R_1^p\frac{1}{g(\vartheta_p)}\right] u'(c_1^p(H, \vartheta_p) \tag{6}$$

$$\lambda f^p(\vartheta_p) = \left[\mu(\vartheta_p) + \pi(\vartheta_p)\frac{g'(\vartheta_p)}{1 - g(\vartheta_p)} - \rho(\vartheta_p)R_1^p\frac{1}{1 - g(\vartheta_p)}\right] u'(c_1^p(L, \vartheta_p)) \tag{7}$$

Equating and multiplying both sides by $(1 - g(\vartheta_p))g(\vartheta_p)$:

$$\frac{(1 - g(\vartheta_p))g(\vartheta_p)}{u'(c_1^p(H, \vartheta_p)}\left[\mu(\vartheta_p) + \pi(\vartheta_p)\frac{g'(\vartheta_p)}{1 - g(\vartheta_p)} - \rho(\vartheta_p)R_1^p\frac{1}{1 - g(\vartheta_p)}\right]$$
$$= \frac{(1 - g(\vartheta_p))g(\vartheta_p)}{u'(c_1^p(L, \vartheta_p)}\left[\mu(\vartheta_p) - \pi(\vartheta_p)\frac{g'(\vartheta_p)}{g(\vartheta_p)} + \rho(\vartheta_p)R_1^p\frac{1}{g(\vartheta_p)}\right]$$

Simplifying:

$$\frac{g(\vartheta_p)}{u'(c_1^p(H, \vartheta_p))}\left[(1 - g(\vartheta_p)) + \pi(\vartheta_p)\frac{g'(\vartheta_p)}{\mu(\vartheta_p)} - \rho(\vartheta_p)R_1^p\frac{1}{\mu(\vartheta_p)}\right]$$
$$= \frac{(1 - g(\vartheta_p))}{u'(c_1^p(L, \vartheta_p))}\left[g(\vartheta_p) - \pi(\vartheta_p)\frac{g'(\vartheta_p)}{g(\vartheta_p)} + \rho(\vartheta_p)R_1^p\frac{1}{g(\vartheta_p)}\right]$$

Let $x = \pi(\vartheta_p)\frac{g'(\vartheta_p)}{\mu(\vartheta_p)} - \rho(\vartheta_p)R_1^p\frac{1}{\mu(\vartheta_p)}$:

$$\frac{(1 - g(\vartheta_p))g(\vartheta_p)}{u'(c_1^p(H, \vartheta_p))} + x\frac{g(\vartheta_p)}{u'(c_1^p(H, \vartheta_p))} = \frac{(1 - g(\vartheta_p))g(\vartheta_p)}{u'(c_1^p(L, \vartheta_p))} - x\frac{1 - g(\vartheta_p)}{u'(c_1^p(L, \vartheta_p))}$$

Rearranging:

$$g(\vartheta_p)\frac{1}{u'(c_1^p(H, \vartheta_p))} + (1 - g(\vartheta_p))\frac{1}{u'(c_1^p(L, \vartheta_p))} = \frac{g(\vartheta_p)(1 - g(\vartheta_p))}{x}\left[\frac{1}{u'(c_1^p(L, \vartheta_p))} - \frac{1}{u'(c_1^p(H, \vartheta_p))}\right]$$

$$\Rightarrow \mathbb{E}\left(\frac{1}{u'(c_1^p(\sigma, \vartheta_p))}\right) = g(\vartheta_p)(1 - g(\vartheta_p))\left[\frac{\mu(\vartheta_p)}{\rho(\vartheta_p)R_1^p - \pi(\vartheta_p)g'(\vartheta_p)}\right]\left[\frac{1}{u'(c_1^p(H, \vartheta_p))} - \frac{1}{u'(c_1^p(L, \vartheta_p))}\right]$$

From FOC (2):

$$\lambda f^p(\vartheta_p) = \mu(\vartheta_p)u'(c_0^p(\vartheta_p))$$

Hence, (6) and (7) can be rewritten:

$$u'(c_0^p(\vartheta_p)) = \left[1 - \frac{\pi(\vartheta_p)}{\mu(\vartheta_p)}\frac{g'(\vartheta_p)}{g(\vartheta_p)} + \frac{\rho(\vartheta_p)}{\mu(\vartheta_p)}R_1^p\frac{1}{g(\vartheta_p)}\right] u'(c_1^p(H, \vartheta_p) \qquad (6)$$

$$\Rightarrow \frac{g(\vartheta_p)(1-g(\vartheta_p))}{u'(c_1^p(H, \vartheta_p)} = \left[g(\vartheta_p) - \pi(\vartheta_p)\frac{g'(\vartheta_p)}{\mu(\vartheta_p)} + \frac{\rho(\vartheta_p)}{\mu(\vartheta_p)}R_1^p\right]\frac{1-g(\vartheta_p)}{u'(c_0^p(\vartheta_p))} \qquad (6i)$$

$$u'(c_0^p(\vartheta_p)) = \left[1 + \pi(\vartheta_p)\frac{g'(\vartheta_p)}{1-g(\vartheta_p)} - \rho(\vartheta_p)R_1^p\frac{1}{1-g(\vartheta_p)}\right] u'(c_1^p(L, \vartheta_p)) \qquad (7)$$

$$\Rightarrow \frac{g(\vartheta_p)(1-g(\vartheta_p))}{u'(c_1^p(L, \vartheta_p)} = \left[(1-g(\vartheta_p)) + \pi(\vartheta_p)\frac{g'(\vartheta_p)}{\mu(\vartheta_p)} - \frac{\rho(\vartheta_p)}{\mu(\vartheta_p)}R_1^p\right]\frac{g(\vartheta_p)}{u'(c_0^p(\vartheta_p))} \qquad (7i)$$

Subtracting (7i) from (6i) and simplifying:

$$g(\vartheta_p)(1-g(\vartheta_p))\left[\frac{1}{u'(c_1^p(H, \vartheta_p))} - \frac{1}{u'(c_1^p(L, \vartheta_p))}\right]$$
$$= \frac{1-g(\vartheta_p)}{u'(c_0^p(\vartheta_p))}\left[g(\vartheta_p) + \frac{\rho(\vartheta_p)R_1^p}{\mu(\vartheta_p)} - \frac{\pi(\vartheta_p)g'(\vartheta_p)}{\mu(\vartheta_p)}\right] - \frac{g(\vartheta_p)}{u'(c_1^p(\vartheta_p))}\left[(1-g(\vartheta_p)) - \frac{\rho(\vartheta_p)R_1^p}{\mu(\vartheta_p)} + \frac{\pi(\vartheta_p)g'(\vartheta_p)}{\mu(\vartheta_p)}\right]$$

Let $y = g(\vartheta_p) + \frac{\rho(\vartheta_p)R_1^p}{\mu(\vartheta_p)} - \frac{\pi(\vartheta_p)g'(\vartheta_p)}{\mu(\vartheta_p)}$, then:

$$g(\vartheta_p)(1-g(\vartheta_p))\left[\frac{1}{u'(c_1^p(H, \vartheta_p))} - \frac{1}{u'(c_1^p(L, \vartheta_p))}\right] = \frac{1}{u'(c_0^p(\vartheta_p))}\Big[y(1-g(\vartheta_p)) - (1-y)g(\vartheta_p)\Big]$$

$$\Rightarrow g(\vartheta_p)(1-g(\vartheta_p))\left[\frac{1}{u'(c_1^p(H, \vartheta_p))} - \frac{1}{u'(c_1^p(L, \vartheta_p))}\right] = \frac{1}{u'(c_0^p(\vartheta_p))}\Big[y - g(\vartheta_p)\Big]$$

Hence:

$$g(\vartheta_p)(1-g(\vartheta_p))\left[\frac{1}{u'(c_1^p(H, \vartheta_p))} - \frac{1}{u'(c_1^p(L, \vartheta_p))}\right] = \frac{1}{u'(c_0^p(\vartheta_p))}\Big[\frac{\rho(\vartheta_p)R_1^p - \pi(\vartheta_p)g'(\vartheta_p)}{\mu(\vartheta_p)}\Big]$$

$$\frac{1}{u'(c_0^p(\vartheta_p))} = g(\vartheta_p)(1-g(\vartheta_p))\left[\frac{\mu(\vartheta_p)}{\rho(\vartheta_p)R_1^p - \pi(\vartheta_p)g'(\vartheta_p)}\right]\left[\frac{1}{u'(c_1^p(H, \vartheta_p))} - \frac{1}{u'(c_1^p(L, \vartheta_p))}\right]$$

Therefore:

$$\frac{1}{u'(c_0^p(\vartheta_p))} = \mathbb{E}_\sigma\left(\frac{1}{u'(c_1^p(\sigma, \vartheta_p))}\right)$$

Note: Since $\frac{1}{u'(c_1^p(\vartheta_p))}$ is convex, by Jensen's inequality;

$$\mathbb{E}_\sigma\left(\frac{1}{u'(c_1^p(\vartheta_p))}\right) \geq \frac{1}{\mathbb{E}_\sigma(u'(c_1^p(\vartheta_p)))}$$

Intertemporal Wedge: Child (High Type) $t = 1$ **and** $t = 2$**:**

$$\frac{\lambda g(\vartheta_p) f^p(\vartheta_p)}{R_1^p} = \omega(\vartheta_p, H) u'(c_1^k(H, \vartheta_p)) \tag{10}$$

$$\frac{\lambda g(\vartheta_p) f^p(\vartheta_p)}{R_1^p} = \frac{R_2^k(1)}{f^k(\vartheta_k|H,1)} \varphi(H, \vartheta_p, \vartheta_k) u'(c_2^k(H, \vartheta_k, \vartheta_p)) \tag{12}$$

$$\frac{\partial \psi(H, \vartheta_p, \vartheta_k)}{\partial \vartheta_k} = \beta_2^1 \omega(H, \vartheta_p) f^k(\vartheta_k|H,1) - \varphi(H, \vartheta_p, \vartheta_k) \tag{16}$$

From (16):

$$\psi(H, \vartheta_p, \vartheta_k) = \beta_2^1 \omega(H, \vartheta_p) \int_{\underline{\vartheta}_k}^{\vartheta_k} f^k(x|H,1) dx - \int_{\underline{\vartheta}_k}^{\vartheta_k} \varphi(H, \vartheta_p, x) dx \tag{A.18}$$

$$\int_{\underline{\vartheta}_k}^{\vartheta_k} \frac{\varphi(H, \vartheta_p, x)}{\beta_2^1 \omega(H, \vartheta_p)} dx = \int_{\underline{\vartheta}_k}^{\vartheta_k} f^k(x|H,1) dx - \frac{\psi(H, \vartheta_p, \vartheta_k)}{\beta_2^1 \omega(H, \vartheta_p)}$$

Therefore, by substitution:

$$\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{\varphi(H, \vartheta_p, \vartheta_k)}{\beta_2^1 \omega(H, \vartheta_p)} d\vartheta_k = \int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} f^k(\vartheta_k|H,1) dx - \frac{\psi(H, \vartheta_p, \overline{\vartheta}_k)}{\beta_2^1 \omega(H, \vartheta_p)}$$

However, $\frac{\psi(H, \vartheta_p, \overline{\vartheta}_k)}{\beta_2^1 \omega(H, \vartheta_p)} = 0$ (terminal condition), hence:

$$\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{\varphi(H, \vartheta_p, \vartheta_k)}{\beta_2^1 \omega(H, \vartheta_p)} d\vartheta_k = F^k(\overline{\vartheta}_k|H,1) - F^k(\underline{\vartheta}_k|H,1)$$

$$\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{\varphi(H, \vartheta_p, \vartheta_k)}{\beta_2^1 \omega(H, \vartheta_p)} d\vartheta_k = 1 \tag{A.19}$$

Equating (10) and (12) and integrating between $\overline{\vartheta}_k$ and $\underline{\vartheta}_k$ wrt ϑ_k:

$$\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{1}{u'(c_2^k(H, \vartheta_p, \vartheta_k))} f^k(\vartheta_k|H,1) d\vartheta_k = \frac{1}{u'(c_1^k(H, \vartheta_p))} \int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{\varphi(H, \vartheta_p, \vartheta_k)}{\omega(H, \vartheta_p) \beta_2^1} d\vartheta_k$$

Using (19):

$$\therefore \frac{1}{u'(c_1^k(H, \vartheta_p))} = \mathbb{E}_{\vartheta_k} \left[\frac{1}{u'(c_2^k(H, \vartheta_p, \vartheta_k))} \right]$$

Intertemporal Wedge: Child (Low Type) $t = 1$ **and** $t = 2$**:**

$$\frac{\lambda(1 - g(\vartheta_p))f^p(\vartheta_p)}{R_1^p} = \left[\omega(L, \vartheta_p, \vartheta_k) - \alpha\rho(\vartheta_p)\right]u'(c_1^k(L, \vartheta_p)) \tag{11}$$

$$\frac{\lambda(1 - g(\vartheta_p))f^p(\vartheta_p)}{R_1^p} = \frac{R_2^k(0)}{f^k(\vartheta_k|L, 0)}\left[\varphi(L, \vartheta_p, \vartheta_k) - \alpha\rho(\vartheta_p)\beta_2^0 f^k(\vartheta_k|H, 0)\right]u'(c_2^k(L, \vartheta_p, \vartheta_k)) \tag{13}$$

$$\frac{\partial\psi(L, \vartheta_p, \vartheta_k)}{\partial\vartheta_k} = \beta_2^0\omega(L, \vartheta_p)f^k(\vartheta_k|L, 0) - \varphi(L, \vartheta_p, \vartheta_k) \tag{16}$$

By symmetry of the expression (19), (16) can be expressed as:

$$\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{\varphi(L, \vartheta_p, \vartheta_k)}{\beta_2^0\omega(L, \vartheta_p)} d\vartheta_k = 1$$

Equating (11) and (13) and integrating between $\overline{\vartheta}_k$ and $\underline{\vartheta}_k$ wrt ϑ_k:

$$\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{1}{u'(c_2^k(L, \vartheta_p, \vartheta_k))} f^k(\vartheta_k|L, 0) d\vartheta_k = \frac{1}{u'(c_1^k(L, \vartheta_p))}\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{R_2^k(0)\left[\varphi(L, \vartheta_p, \vartheta_k) - \alpha\rho(\vartheta_p)\beta_2^0 f^k(\vartheta_k|H, 0)\right]}{\omega(L, \vartheta_p) - \alpha\rho(\vartheta_p)} d\vartheta_k$$

$$\left[\omega(L, \vartheta_p) - \alpha\rho(\vartheta_p)\right]\mathbb{E}_{\vartheta_k}\left[\frac{1}{u'(c_2^k(L, \vartheta_p, \vartheta_k))}\right] = \left[\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \left[R_2^k(0)\varphi(L, \vartheta_p, \vartheta_k)\right]d\vartheta_k - \alpha\rho(\vartheta_p)\right]\frac{1}{u'(c_1^k(L, \vartheta_p))}$$

Dividing both sides by $\omega(L, \vartheta_p)$ and since $R_2^k(0) = \frac{1}{\beta_2^0}$:

$$\left[1 - \alpha\rho(\vartheta_p)\right]\mathbb{E}_{\vartheta_k}\left[\frac{1}{u'(c_2^k(L, \vartheta_p, \vartheta_k))}\right] = \left[\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{\varphi(L, \vartheta_p, \vartheta_k)}{\beta_2^0\omega(L, \vartheta_p)} d\vartheta_k - \alpha\rho(\vartheta_p)\right]\frac{1}{u'(c_1^k(L, \vartheta_p))}$$

$$\left[1 - \alpha\rho(\vartheta_p)\right]\mathbb{E}_{\vartheta_k}\left[\frac{1}{u'(c_2^k(L, \vartheta_p, \vartheta_k))}\right] = \left[1 - \alpha\rho(\vartheta_p)\right]\frac{1}{u'(c_1^k(L, \vartheta_p))}$$

$$\therefore \frac{1}{u'(c_1^k(L, \vartheta_p))} = \mathbb{E}_{\vartheta_k}\left[\frac{1}{u'(c_2^k(L, \vartheta_p, \vartheta_k))}\right]$$

A.0.4 Proof of Proposition 2

Labour Wedge $t = 0$**: Parent**:

From FOC (1):

$$\frac{\partial \pi(\vartheta_p)}{\partial \vartheta_p} = \gamma f^p(\vartheta_p) - \mu(\vartheta_p)$$

$$\pi(\vartheta_p) = \int_{\vartheta_p}^{\overline{\vartheta}_p} \left(\mu(x) - \gamma f^p(x)\right) dx \tag{1i}$$

From FOC (2):

$$\mu(\vartheta_p) = \frac{\lambda f^p(\vartheta_p)}{u'(c_0^p(\vartheta_p))} \tag{2i}$$

Hence:

$$\pi(\vartheta_p) = \int_{\vartheta_p}^{\overline{\vartheta}_p} \left(\frac{\lambda f^p(x)}{u'(c_0^p(x))} - \gamma f^p(x) \right) dx \tag{A.20}$$

Substituting (1i) and (2i) into FOC (3):

$$\lambda f^p(\vartheta_p) - \frac{\lambda f^p(\vartheta_p)}{u'(c_0^p(\vartheta_p))} \left[\frac{1}{\vartheta_p} h' \left(\frac{y_0^p(\vartheta_p)}{\vartheta_p} \right) \right]$$
$$= \int_{\vartheta_p}^{\overline{\vartheta}_p} \left(\frac{\lambda f^p(x)}{u'(c_0^p(x))} - \gamma f^p(x) \right) dx \left[\frac{1}{\vartheta_p^2} h' \left(\frac{y_0^p(\vartheta_p)}{\vartheta_p} \right) + \frac{y_0^p(\vartheta_p)}{\vartheta_p^3} h'' \left(\frac{y_0^p(\vartheta_p)}{\vartheta_p} \right) \right]$$

Dividing by $\frac{1}{\vartheta_p} h' \left(\frac{y_0^p(\vartheta_p)}{\vartheta_p} \right)$ and λ:

$$f^p(\vartheta_p) \left[\frac{1}{\frac{1}{\vartheta_p} h' \left(\frac{y_0^p(\vartheta_p)}{\vartheta_p} \right)} - \frac{1}{u'(c_0^p(\vartheta_p))} \right] = \int_{\vartheta_p}^{\overline{\vartheta}_p} \left(\frac{f^p(x)}{u'(c_0^p(x))} - \frac{\gamma f^p(x)}{\lambda} \right) dx \left[\frac{1}{\vartheta_p} + \frac{y_0^p(\vartheta_p)}{\vartheta_p^2} \frac{h'' \left(\frac{y_0^p(\vartheta_p)}{\vartheta_p} \right)}{h' \left(\frac{y_0^p(\vartheta_p)}{\vartheta_p} \right)} \right]$$

Define: $\tau_0^w(\vartheta_p) = 1 - \frac{h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)}{\vartheta_p u'(c_0^p(\vartheta_p))}$, then:

$$\frac{1}{u'(c_0^p(\vartheta_p))(1 - \tau_0^w(\vartheta_p))} - \frac{1}{u'(c_0^p(\vartheta_p))} = \frac{1}{f^p(\vartheta_p)} \left[\frac{1}{\vartheta_p} + \frac{y_0^p(\vartheta_p)}{\vartheta_p^2} \frac{h'' \left(\frac{y_0^p(\vartheta_p)}{\vartheta_p} \right)}{h' \left(\frac{y_0^p(\vartheta_p)}{\vartheta_p} \right)} \right]$$
$$\int_{\vartheta_p}^{\overline{\vartheta}_p} \left(\frac{f^p(x)}{u'(c_0^p(x))} - \frac{\gamma f^p(x)}{\lambda} \right) dx$$

Multiplying by $\frac{1-F^p(\vartheta_p)}{1-F^p(\vartheta_p)}$ and simplifying:

$$\frac{\tau_0^w(\vartheta_p)}{1-\tau_0^w(\vartheta_p)} = \frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left[1+\frac{y_0^p(\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_0^p(\vartheta_p)}{\vartheta_p}\right)}\right]\left[\int_{\vartheta_p}^{\overline{\vartheta}_p}\frac{u'(c_0^p(\vartheta_p))}{u'(c_0^p(x))}\left(1-\frac{\gamma u'(c_0^p(x))}{\lambda}\right)\frac{f^p(x)}{1-F^p(\vartheta_p)}dx\right]$$

Notice: $\pi(\underline{\vartheta}_p) = \int_{\underline{\vartheta}_p}^{\overline{\vartheta}_p}\mu(\vartheta_p)d\vartheta_p - \gamma$, but $\pi(\underline{\vartheta}_p) = 0$ (Terminal Condition). Therefore:

$$\gamma = \int_{\underline{\vartheta}_p}^{\overline{\vartheta}_p}\frac{\lambda}{u'(c_0^p(\vartheta_p))}f^p(\vartheta_p)d\vartheta_p$$

$$\frac{1}{\lambda} = \mathbb{E}_{\vartheta_p}\left[\frac{1}{\gamma u'(c_0^p(\vartheta_p))}\right]$$

Hence the term $\left(1-\frac{\gamma u'(c_0^p(x))}{\lambda}\right)$ can be rewritten as:

$$\left(1-\frac{\gamma u'(c_0^p(x))}{\lambda}\right) = 1-\frac{\mathbb{E}_{\vartheta_p}\left[\frac{1}{u'(c_0^p(\vartheta_p))}\right]}{\frac{1}{u'(c_0^p(x))}}$$

Labour Wedge $t=1$: Parent Send Child to College:
From FOC (6) and (8):

$$\frac{\lambda g(\vartheta_p)f^p(\vartheta_p)}{R_1^p} = \left[\mu(\vartheta_p)\beta g(\vartheta_p) - \pi(\vartheta_p)\beta g'(\vartheta_p) + \rho(\vartheta_p)\right]u'(c_1^p(H,\vartheta_p)) \tag{6}$$

$$\begin{aligned}\frac{\lambda g(\vartheta_p)f^p(\vartheta_p)}{R_1^p} &= \left[\mu(\vartheta_p)\beta g(\vartheta_p) - \pi(\vartheta_p)\beta g'(\vartheta_p) + \rho(\vartheta_p)\right]\frac{1}{\vartheta_p}h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right) + \\ &\quad \pi(\vartheta_p)\beta g(\vartheta_p)\left(\frac{1}{\vartheta_p^2}h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right) + \frac{y_1^p(H,\vartheta_p)}{\vartheta_p^3 h''\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}\right)\end{aligned} \tag{8}$$

Using (6), equation (8) becomes:

$$\frac{\lambda g(\vartheta_p) f^p(\vartheta_p)}{R_1^p}\left[1 - \frac{1}{u'(c_1^p(H,\vartheta_p))}\frac{1}{\vartheta_p}h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)\right] =$$
$$\beta g(\vartheta_p)\left[\frac{1}{\vartheta_p^2}h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right) + \frac{y_1^p(H,\vartheta_p)}{\vartheta_p^3}h''\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)\right]\pi(\vartheta_p)$$

Dividing by $\frac{\lambda g(\vartheta_p) f^p(\vartheta_p)}{R_1^p}\frac{1}{\vartheta_p}h'\left(\frac{y_1^k(H,\vartheta_p)}{\vartheta_p}\right)$ and simplifying:

$$\left[\frac{1}{\frac{1}{\vartheta_p}h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)} - \frac{1}{u'(c_1^p(H,\vartheta_p))}\right] = \pi(\vartheta_p)\frac{1}{\lambda\vartheta_p f^p(\vartheta_p)}\left[1 + \frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}\right]$$

Define $\tau_1^w(H,\vartheta_p) = 1 - \frac{h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}{\vartheta_p u'(c_1^p(H,\vartheta_p))}$:

$$\frac{\tau_1^w(H,\vartheta_p)}{1-\tau_1^w(H,\vartheta_p)} = \pi(\vartheta_p)\frac{u'(c_1^p(H,\vartheta_p))}{\lambda\vartheta_p f^p(\vartheta_p)}\left[1 + \frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}\right]$$

From FOC (1), (2), and (6):

$$\pi(\vartheta_p) = \int_{\vartheta_p}^{\overline{\vartheta}_p}\Big(\mu(x) - \gamma f^p(x)\Big)dx \tag{1}$$

$$\mu(x) = \frac{\lambda f^p(x)}{u'(c_0^p(x))} \tag{2}$$

$$\mu(x) = \frac{\lambda f^p(x)}{u'(c_1^p(H,x))} + \frac{\pi(x)g'(x)}{g(x)} - \frac{\rho(x)}{\beta g(x)} \tag{6}$$

To get the labour wedge in terms of the parent's consumption choices, we use (1) and (2):

$$\pi(\vartheta_p) = \int_{(\vartheta_p)}^{\overline{\vartheta}_p}\left(\frac{\lambda f^p(x)}{u'(c_0^p(x))} - \gamma f^p(x)\right)dx$$

Adding and subtracting $\frac{\lambda f^p(x)}{u'(c_1^p(x))}$ and simplifying:

$$\pi(\vartheta_p) = \int_{(\vartheta_p)}^{\overline{\vartheta}_p}\left(\frac{1}{u'(c_1^p(x))} - \frac{\gamma}{\lambda}\right)\frac{f^p(x)}{\lambda}dx - \int_{(\vartheta_p)}^{\overline{\vartheta}_p}\left(\frac{1}{u'(c_1^p(x))} - \frac{1}{u'(c_0^p(x))}\right)\frac{f^p(x)}{\lambda}dx$$

Substituting the above into the labour wedge, and multiplying by $\frac{1-F^{p}(\vartheta_p)}{1-F^{p}(\vartheta_p)}$:

$$\frac{\tau_1^w(H,\vartheta_p)}{1-\tau_1^w(H,\vartheta_p)} =$$

$$\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}\right)\int_{\vartheta_p}^{\overline{\vartheta}_p}\frac{u'(c_1^p(H,\vartheta_p))}{u'(c_1^p(H,x))}\left(1-\frac{\gamma u'(c_1^p(H,\vartheta_p))}{\lambda}\right)\frac{f^p(x)}{1-F(\vartheta_p)}dx$$

$$-\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}\right)\int_{\vartheta_p}^{\overline{\vartheta}_p}\frac{u'(c_1^p(H,\vartheta_p))}{u'(c_1^p(H,x))}\left[1-\frac{u'(c_1^p(H,x))}{u'(c_0^p(x))}\right]\frac{f^p(x)}{1-F^p(\vartheta_p)}dx$$

Alternatively we can find an expression for the labour wedge in terms of the parent and child's consumption. Substituting (6) into (1), and multiplying by $\frac{1-F^{p}(\vartheta_p)}{1-F^{p}(\vartheta_p)}$ we get:

$$\frac{\tau_1^w(H,\vartheta_p)}{1-\tau_1^w(H,\vartheta_p)} =$$

$$\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}\right)\int_{\vartheta_p}^{\overline{\vartheta}_p}\frac{u'(c_1^p(H,\vartheta_p))}{u'(c_1^p(H,x))}\left(1-\frac{\gamma u'(c_1^p(H,x))}{\lambda}\right)\frac{f^p(x)}{1-F(\vartheta_p)}dx$$

$$-\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}\right)\left[\int_{\vartheta_p}^{\overline{\vartheta}_p}\frac{\rho(x)-\beta\pi(x)g'(x)}{\beta g(x)}\frac{u'(c_1^p(H,\vartheta_p))}{\lambda(1-F^p(\vartheta_p))}dx\right]$$

From FOCs (4) and (10):

$$\frac{\rho(x)-\beta\pi(x)g'(x)}{\beta g(x)} = \frac{1}{\alpha}\left[\frac{\omega(H,x)}{\beta g(x)}-(1-\gamma)f^p(x)-\alpha\mu(x)\right] \tag{4}$$

$$\omega(H,x) = \frac{\lambda g(x) f^p(x)}{R_1^p u'(c_1^k(H,x))} \tag{10}$$

substituting (10) and (2) into (4), and then subtituting (4) back into the labour wedge:

$$\frac{\tau_1^w(H,\vartheta_p)}{1-\tau_1^w(H,\vartheta_p)} =$$

$$\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}\right)\int_{\vartheta_p}^{\bar{\vartheta}_p}\frac{u'(c_1^p(H,\vartheta_p))}{u'(c_1^p(H,x))}\left(1-\frac{\gamma u'(c_1^p(H,x))}{\lambda}\right)\frac{f^p(x)}{1-F(\vartheta_p)}dx$$

$$-\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(H,\vartheta_p)}{\vartheta_p}\right)}\right)\int_{\vartheta_p}^{\bar{\vartheta}_p}\left[\left(\frac{u'(c_1^p(H,\vartheta_p))}{\alpha u'(c_1^k(H,x))}-\frac{u'(c_1^p(H,\vartheta_p))}{u'(c_0^p(x))}\right)\right.$$
$$\left.-\frac{(1-\gamma)u'(c_1^p(H,\vartheta_p))}{\alpha\lambda}\right]\frac{f^p(x)}{1-F^p(\vartheta_p)}dx$$

Given the two expressions for the (H,ϑ_p) labour wedge at $t=1$ implies:

$$\frac{u'(c_1^p(H,\vartheta_p))}{u'(c_1^p(H,x))}\left[1-\frac{u'(c_1^p(H,x))}{u'(c_0^p(x))}\right]=\left(\frac{u'(c_1^p(H,\vartheta_p))}{\alpha u'(c_1^k(H,x))}-\frac{u'(c_1^p(H,\vartheta_p))}{u'(c_0^p(x))}\right)-\frac{(1-\gamma)u'(c_1^p(H,\vartheta_p))}{\alpha\lambda}$$

$$\frac{u'(c_1^p(H,\vartheta_p))}{u'(c_1^p(H,x))}\left[1-\frac{u'(c_1^p(H,x))}{u'(c_0^p(x))}\right]=\frac{u'(c_1^p(H,\vartheta_p))}{u'(c_1^p(H,x))}\left(\frac{u'(c_1^p(H,x))}{\alpha u'(c_1^k(H,x))}-\frac{u'(c_1^p(H,x))}{u'(c_0^p(x))}-\frac{(1-\gamma)u'(c_1^p(H,x))}{\alpha\lambda}\right)$$

Hence:

$$1-\frac{u'(c_1^p(H,x))}{u'(c_0^p(x))}=\frac{u'(c_1^p(H,x))}{\alpha u'(c_1^k(H,x))}-\frac{u'(c_1^p(H,x))}{u'(c_0^p(x))}-\frac{(1-\gamma)u'(c_1^p(H,x))}{\alpha\lambda}$$

Labour Wedge $t=1$: Parent Does Not Send Child to College:
Using FOC (7) and (9), and following the same steps as for high type leads to a symmetric expression:

$$\frac{\tau_1^w(L,\vartheta_p)}{1-\tau_1^w(l,\vartheta_p)}=\pi(\vartheta_p)+\frac{u'(c_1^p(L,\vartheta_p))}{\vartheta_p\lambda f^p(\vartheta_p)}\left[1+\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}\right]$$

Hence the labour wedge at $t=1$ when the child is low innate ability is symmetric to that of a high type:

$$\frac{\tau_1^w(L,\vartheta_p)}{1-\tau_1^w(L,\vartheta_p)} =$$

$$\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}\right)\int_{\vartheta_p}^{\overline{\vartheta}_p}\frac{u'(c_1^p(L,\vartheta_p))}{u'(c_1^p(L,x))}\left(1-\frac{\gamma u'(c_1^p(L,\vartheta_p))}{\lambda}\right)\frac{f^p(x)}{1-F(\vartheta_p)}dx$$

$$-\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}\right)\int_{\vartheta_p}^{\overline{\vartheta}_p}\frac{u'(c_1^p(L,\vartheta_p))}{u'(c_1^p(L,x))}\left[1-\frac{u'(c_1^p(L,x))}{u'(c_0^p(x))}\right]\frac{f^p(x)}{1-F^p(\vartheta_p)}dx$$

Alternatively, if we can derive the labour wedge in terms of the parent and child's consumption as well. From FOC (1) and (7):

$$\pi(\vartheta_p) = \int_{\vartheta_p}^{\overline{\vartheta}_p}\Big(\mu(x)-\gamma f^p(x)\Big)dx \tag{1}$$

$$\mu(x) = \frac{\lambda f^p(x)}{u'(c_1^p(L,x))} - \frac{\pi(x)g'(x)}{1-g(x)} + \frac{\rho(x)}{\beta(1-g(x))} \tag{7}$$

Substituting (7) into (1):

$$\pi(\vartheta_p) = \int_{\vartheta_p}^{\overline{\vartheta}_p}\left(\frac{\lambda f^p(x)}{u'(c_1^p(L,x))} - \gamma f^p(x) - \frac{\pi(x)g'(x)\beta-\rho(x)}{\beta(1-g(x))}\right)dx$$

Therefore, the labour wedge can be expressed as:

$$\frac{\tau_1^w(L,\vartheta_p)}{1-\tau_1^w(L,\vartheta_p)} =$$

$$\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}\right)\int_{\vartheta_p}^{\overline{\vartheta}_p}\frac{u'(c_1^p(L,\vartheta_p))}{u'(c_1^p(L,x))}\left(1-\frac{\gamma u'(c_1^p(L,x))}{\lambda}\right)\frac{f^p(x)}{1-F(\vartheta_p)}dx$$

$$-\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}\right)\left[\int_{\vartheta_p}^{\overline{\vartheta}_p}\frac{\beta\pi(x)g'(x)-\rho(x)}{\beta(1-g(x))}\frac{u'(c_1^p(L,\vartheta_p))}{\lambda(1-F^p(\vartheta_p))}dx\right]$$

Using FOC (5) and (11):

$$\frac{\beta\pi(x)g'(x)-\rho(x)}{\beta(1-g(x))} = \frac{1}{\alpha}\left[\frac{\lambda f^p(x)}{u'(c_1^k(L,x))} - (1-\gamma)f^p(x)\right] - \mu(x)$$

From here, we again see that the equations are symmetric to the high type. Hence:

$$
\begin{aligned}
&\frac{\tau_1^w(L,\vartheta_p)}{1-\tau_1^w(L,\vartheta_p)} = \\
&\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}\right)\int_{\vartheta_p}^{\overline{\vartheta}_p}\frac{u'(c_1^p(L,\vartheta_p))}{u'(c_1^p(L,x))}\left(1-\frac{\gamma u'(c_1^p(L,\vartheta_p))}{\lambda}\right)\frac{f^p(x)}{1-F(\vartheta_p)}dx \\
&-\frac{1-F^p(\vartheta_p)}{\vartheta_p f^p(\vartheta_p)}\left(1+\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\frac{h''\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}{h'\left(\frac{y_1^p(L,\vartheta_p)}{\vartheta_p}\right)}\right)\int_{\vartheta_p}^{\overline{\vartheta}_p}\left[\lambda\left(\frac{u'(c_1^p(L,\vartheta_p))}{\alpha u'(c_1^k(L,x))}-\frac{u'(c_1^p(L,\vartheta_p))}{u'(c_0^p(x))}\right)\right. \\
&\qquad\qquad\left.-\frac{(1-\gamma)u'(c_1^p(L,\vartheta_p))}{\alpha}\right]\frac{f^p(x)}{\lambda(1-F^p(\vartheta_p))}dx
\end{aligned}
$$

Labour Wedge $t=2$: Child of High Type:

$$
\varphi(H,\vartheta_p,\vartheta_k)=\frac{\lambda g(\vartheta_p)f^p(\vartheta_p)f^k(\vartheta_k|H,1)}{R_1^pR_2^k(1)}\frac{1}{u'(c_2^k(H,\vartheta_p,\vartheta_k))} \tag{12}
$$

$$
\begin{aligned}
&\lambda\frac{g(\vartheta_p)}{R_1^p}\frac{1}{R_2^k(1)}f^k(\vartheta_k|H,1)f^p(\vartheta_p)-\varphi(H,\vartheta_p,\vartheta_k)\frac{1}{\vartheta_k}h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right) \\
&=\psi(H,\vartheta_p,\vartheta_k)\left[\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k^3}h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)+\frac{1}{\vartheta_k^2}h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)\right]
\end{aligned} \tag{14}
$$

$$
\psi(H,\vartheta_p,\vartheta_k)=\int_{\vartheta_k}^{\overline{\vartheta}_k}\left[\varphi(H,\vartheta_p,x)-\beta_2^1\omega(H,\vartheta_p)f^k(x|H,1)\right]dx \tag{21}
$$

Substituting (12) and (18) into (14):

$$
\begin{aligned}
&\frac{\lambda g(\vartheta_p)f^p(\vartheta_p)f^k(\vartheta_k|H,1)}{R_1^pR_2^k(1)}\left[1-\frac{1}{u'(c_2^k(H,\vartheta_p,\vartheta_k))}\frac{1}{\vartheta_k}h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)\right]= \\
&\qquad\int_{\vartheta_k}^{\overline{\vartheta}_k}\left[\frac{\lambda g(\vartheta_p)f^p(\vartheta_p)f^k(x|H,1)}{R_1^pR_2^k(1)u'(c_2^k(H,\vartheta_p,x))}-\omega(H,\vartheta_p)\beta_2^1f^k(x|H,1)\right]dx \\
&\qquad\qquad\left[\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k^3}h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)+\frac{1}{\vartheta_k^2}h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)\right]
\end{aligned}
$$

Dividing by $\frac{\lambda g(\vartheta_p)f^p(\vartheta_p)f^k(\vartheta_k|H,1)}{R_1^pR_2^k(1)}\frac{1}{\vartheta_k}h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)$:

$$\left[\frac{1}{\frac{1}{\vartheta_k}h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}-\frac{1}{u'(c_2^k(H,\vartheta_p,\vartheta_k))}\right]=\frac{1}{\vartheta_k f^k(\vartheta_k|H,1)}\int_{\vartheta_k}^{\overline{\vartheta}_k}\left[\frac{f^k(x|H,1)}{u'(c_2^k(H,\vartheta_p,x))}\right.$$

$$\left.-\frac{\omega(H,\vartheta_p)R_1^p f^k(x|H,1)}{\lambda g(\vartheta_p)f^p(\vartheta_p)}\right]dx\left[1+\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]$$

Define: $\tau_2^w(H,\vartheta_p,\vartheta_k)=1-\frac{h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{\vartheta_k u'(c_2^k(H,\vartheta_p,\vartheta_k))}$:

$$\frac{\tau_2^w(H,\vartheta_p,\vartheta_k)}{1-\tau_2^w(H,\vartheta_p,\vartheta_k)}=\frac{1}{\vartheta_k f^k(\vartheta_k|H,1)}\left[1+\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]$$

$$\left[\int_{\vartheta_k}^{\overline{\vartheta}_k}\left(\frac{u'(c_2^k(H,\vartheta_p,\vartheta_k))}{u'(c_2^k(H,\vartheta_p,x))}f^k(x|H,1)-\frac{f^k(x|H,1)u'(c_2^k(H,\vartheta_p,\vartheta_k))}{\lambda}\frac{\omega(H,\vartheta_p)R_1^p}{g(\vartheta_p)f^p(\vartheta_p)}\right)dx\right]$$

Multiplying by $\frac{1-F^k(\vartheta_k|H,1)}{1-F^k(\vartheta_k|H,1)}$:

$$\frac{\tau_2^w(H,\vartheta_p,\vartheta_k)}{1-\tau_2^w(H,\vartheta_p,\vartheta_k)}=\frac{1-F^k(\vartheta_k|H,1)}{\vartheta_k f^k(\vartheta_k|H,1)}\left[1+\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]$$

$$\left[\int_{\vartheta_k}^{\overline{\vartheta}_k}\left(\frac{u'(c_2^k(H,\vartheta_p,\vartheta_k))}{u'(c_2^k(H,\vartheta_p,x))}\frac{f^k(x|H,1)}{1-F^k(\vartheta_k|H,1)}-\frac{(f^k(x|H,1))u'(c_2^k(H,\vartheta_p,\vartheta_k))}{\lambda(1-F^k(\vartheta_k|H,1)}\frac{\omega(H,\vartheta_p)R_1^p}{g(\vartheta_p)f^p(\vartheta_p)}\right)dx\right]$$

Adding and Substracting $(1-\gamma)\frac{(f^k(x|H,1))u'(c_2^k(H,\vartheta_p,\vartheta_k))}{\lambda(1-F^k(\vartheta_k|H,1)}$:

$$\frac{\tau_2^w(H,\vartheta_p,\vartheta_k)}{1-\tau_2^w(H,\vartheta_p,\vartheta_k)} = \frac{1-F^k(\vartheta_k|H,1)}{\vartheta_k f^k(\vartheta_k|H,1)}\left[1+\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]$$
$$\left[\int_{\vartheta_k}^{\overline{\vartheta}_k}\frac{u'(c_2^k(H,\vartheta_p,\vartheta_k))}{u'(c_2^k(H,\vartheta_p,x))}\left(1-\frac{(1-\gamma)u'(c_2^k(H,x))}{\lambda}\right)\frac{f^k(x|H,1)}{1-F^k(\vartheta_k|H,1)}dx\right] - \frac{1-F^k(\vartheta_k|H,1)}{\vartheta_k f^k(\vartheta_k|H,1)}$$
$$\left[1+\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]\left[\int_{\vartheta_k}^{\overline{\vartheta}_k}\left(\frac{\omega(H,\vartheta_p)R_1^p}{g(\vartheta_p)f^p(\vartheta_p)}-(1-\gamma)\right)\frac{(f^k(x|H,1))u'(c_2^k(H,\vartheta_p,\vartheta_k))}{\lambda(1-F^k(\vartheta_k|H,1)}dx\right] \tag{i}$$

From FOCs (4) and (6):

$$\frac{\omega(H,\vartheta_p)R_1^p}{g(\vartheta_p)f^p(\vartheta_p)}-(1-\gamma) = \frac{\alpha}{\beta f^p(\vartheta_p)g(\vartheta_p)}\left[\beta g(\vartheta_p)\mu(\vartheta_p)-\beta g'(\vartheta_p)\pi(\vartheta_p)+\rho(\vartheta_p)\right] \tag{4}$$

$$\beta g(\vartheta_p)\mu(\vartheta_p)-\beta g'(\vartheta_p)\pi(\vartheta_p)+\rho(\vartheta_p) = \frac{\lambda g(\vartheta_p)f^p(\vartheta_p)}{R_1^p u'(c_1^p(H,\vartheta_p))} \tag{6}$$

Substituting these terms into the labour wedge yields:

$$\frac{\tau_2^w(H,\vartheta_p,\vartheta_k)}{1-\tau_2^w(H,\vartheta_p,\vartheta_k)} = \frac{1-F^k(\vartheta_k|H,1)}{\vartheta_k f^k(\vartheta_k|H,1)}\left[1+\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]$$
$$\left[\int_{\vartheta_k}^{\overline{\vartheta}_k}\frac{u'(c_2^k(H,\vartheta_p,\vartheta_k))}{u'(c_2^k(H,\vartheta_p,x))}\left(1-\frac{(1-\gamma)u'(c_2^k(H,x))}{\lambda}\right)\frac{f^k(x|H,1)}{1-F^k(\vartheta_k|H,1)}dx\right]$$
$$-\frac{1-F^k(\vartheta_k|H,1)}{\vartheta_k f^k(\vartheta_k|H,1)}\left[1+\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]\left[\frac{\alpha u'(c_2^k(H,\vartheta_p,\vartheta_k))}{u'(c_1^p(H,\vartheta_p))}\right]$$

Alternative, using only the Child's FOC (10):

$$\frac{\omega(H,\vartheta_p)R_1^p}{g(\vartheta_p)f^p(\vartheta_p)} = \frac{\lambda}{u'(c_1^k(H,\vartheta_p))}$$

The labour wedge equation (i) simplifies to:

$$\frac{\tau_2^w(H,\vartheta_p,\vartheta_k)}{1-\tau_2^w(H,\vartheta_p,\vartheta_k)} = \frac{1-F^k(\vartheta_k|H,1)}{\vartheta_k f^k(\vartheta_k|H,1)}\left[1+\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]$$

$$\left[\int_{\vartheta_k}^{\overline{\vartheta}_k}\frac{u'(c_2^k(H,\vartheta_p,\vartheta_k))}{u'(c_2^k(H,\vartheta_p,x))}\left(1-\frac{(1-\gamma)u'(c_2^k(H,x))}{\lambda}\right)\frac{f^k(x|H,1)}{1-F^k(\vartheta_k|H,1)}dx\right]$$

$$-\frac{1-F^k(\vartheta_k|H,1)}{\vartheta_k f^k(\vartheta_k|H,1)}\left[1+\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(H,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]\left[u'(c_2^k(H,\vartheta_p,\vartheta_k))\left(\frac{1}{u'(c_1^k(H,\vartheta_p))}-\frac{1-\gamma}{\lambda}\right)\right]$$

Labour Wedge $t=2$**: Child of Low Type:**

Define:

$$\tau_2^w(L,\vartheta_p,\vartheta_k) = 1-\frac{h'\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{\vartheta_k u'(c_2^k(L,\vartheta_p,\vartheta_k))}$$

Similar steps as the high type, using FOCs (13), (15) , and (17):

$$\frac{\tau_2^w(L,\vartheta_p,\vartheta_k)}{1-\tau_2^w(L,\vartheta_p,\vartheta_k)} = \frac{1-F^k(\vartheta_k|L,0)}{\vartheta_k f^k(\vartheta_k|L,0)}\left[1+\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]$$

$$\left[\int_{\vartheta_k}^{\overline{\vartheta}_k}\frac{u'(c_2^k(L,\vartheta_p,\vartheta_k))}{u'(c_2^k(L,\vartheta_p,x))}\frac{f^k(x|L,0)}{1-F^k(\vartheta_k|L,0)}dx-\frac{(1-F^k(\vartheta_k|L,0)u'(c_2^k(L,\vartheta_p,\vartheta_k))}{\lambda(1-F^k(\vartheta_k|L,0)}\frac{\omega(L,\vartheta_p)R_1^p}{(1-g(\vartheta_p))f^p(\vartheta_p)}\right.$$
$$\left.+\frac{(1-F^k(\vartheta_k|H,0)u'(c_2^k(L,\vartheta_p,\vartheta_k))}{\lambda(1-F^k(\vartheta_k|L,0)}\frac{\alpha\rho(\vartheta_p)R_1^p}{(1-g(\vartheta_p))f^p(\vartheta_p)}\right]$$

Adding and Substracting $(1-\gamma)\frac{(f^k(x|L,0))u'(c_2^k(L,\vartheta_p,\vartheta_k))}{\lambda(1-F^k(\vartheta_k|L,0)}$, and following similar steps as for high type using FOC (11):

$$\frac{\tau_2^w(L,\vartheta_p,\vartheta_k)}{1-\tau_2^w(L,\vartheta_p,\vartheta_k)} = \frac{1-F^k(\vartheta_k|L,0)}{\vartheta_k f^k(\vartheta_k|L,0)}\left[1+\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]$$

$$\left[\int_{\vartheta_k}^{\overline{\vartheta}_k}\frac{u'(c_2^k(L,\vartheta_p,\vartheta_k))}{u'(c_2^k(L,\vartheta_p,x))}\left(1-\frac{(1-\gamma)u'(c_2^k(L,\vartheta_p,x))}{\lambda}\right)\frac{f^k(x|L,0)}{1-F^k(\vartheta_k|L,0)}dx\right]$$

$$-\frac{1-F^k(\vartheta_k|L,0)}{\vartheta_k f^k(\vartheta_k|L,0)}\left[1+\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]\left[u'(c_2^k(L,\vartheta_p,\vartheta_k))\left(\frac{1}{u'(c_1^k(L,\vartheta_p))}-\frac{(1-\gamma)}{\lambda}\right)\right]$$

$$+\frac{1-F^k(\vartheta_k|L,0)}{\vartheta_k f^k(\vartheta_k|L,0)}\left[1+\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\frac{h''\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}{h'\left(\frac{y_2^k(L,\vartheta_p,\vartheta_k)}{\vartheta_k}\right)}\right]\left[\frac{\alpha\rho(\vartheta_p)R_1^p}{(1-g(\vartheta_p))f^p(\vartheta_p)}\frac{u'(c_2^k(L,\vartheta_p))}{\lambda}\left[1+\frac{1-F^k(\vartheta_k|H,0)}{1-F^k(\vartheta_k|L,0)}\right]\right]$$

A.0.5 Proof of Proposition 4

Intergenerational Wedge $t=0$**:**
Adding FOCs (4) and (5):

$$\omega(H,\vartheta_p)+\omega(L,\vartheta_p)=(1-\gamma)\beta f^p(\vartheta_p)+\mu(\vartheta_p)\alpha\beta+\alpha\rho(\vartheta_p) \tag{A.5.1}$$

From FOC (10) and (11)

$$\omega(H,\vartheta_p)=\frac{\lambda g(\vartheta_p)f^p(\vartheta_p)}{R_1^p(\vartheta_p)u'(c_1^k(H,\vartheta_p))} \tag{10}$$

$$\omega(L,\vartheta_p)=\frac{\lambda(1-g(\vartheta_p))f^p(\vartheta_p)}{R_1^p(\vartheta_p)u'(c_1^k(L,\vartheta_p))}+\alpha\rho(\vartheta_p) \tag{11}$$

Substituting into (A.5.1) and simplifying:

$$\left[\frac{g(\vartheta_p)}{u'(c_1^k(H,\vartheta_p))}+\frac{(1-g(\vartheta_p))}{u'(c_1^k(L,\vartheta_p))}\right]=\frac{(1-\gamma)}{\lambda}+\alpha\frac{\mu(\vartheta_p)}{\lambda f^p(\vartheta_p)}$$

From FOC (2), $\frac{\mu(\vartheta_p)}{\lambda f^p(\vartheta_p)}=\frac{1}{u'(c_0^p(\vartheta_p))}$. Therefore:

$$\frac{1}{u'(c_0^p(\vartheta_p))}=\mathbb{E}_\sigma\left[\frac{1}{\alpha u'(c_1^k(\sigma,\vartheta_p))}\right]-\frac{1-\gamma}{\alpha\lambda}$$

Intergenerational Wedge $t=1$**: High Type**
From FOCs (6), (12):

$$\frac{\lambda g(\vartheta_p) f^p(\vartheta_p)}{R_1^p} = \Big[\beta g(\vartheta_p)\mu(\vartheta_p) - \beta g'(\vartheta_p)\pi(\vartheta_p) + \rho(\vartheta_p)\Big] u'(c_1^p(H, \vartheta_p))$$

$$\frac{\lambda g(\vartheta_p) f^p(\vartheta_p)}{R_1^p} = \frac{R_2^k(1)\varphi(H, \vartheta_p, \vartheta_k)}{f^k(\vartheta_k|H, 1)} u'(c_2^k(H, \vartheta_p, \vartheta_k))$$

Equating and using FOC (4)

$$\Big[\omega(H, \vartheta_p) - (1-\gamma)\beta g(\vartheta_p) f^p(\vartheta_p)\Big] \frac{1}{\alpha} \frac{f^k(\vartheta_k|H, 1)}{u'(c_2^k(H, \vartheta_p, \vartheta_k))} = \frac{\varphi(H, \vartheta_p, \vartheta_k) R_2^k(1)}{u'(c_1^p(H, \vartheta_p))}$$

Integrating both sides wrt ϑ_k from upper bound $\overline{\vartheta}_k$ to lower bound $\underline{\vartheta}_k$

$$\Big[\omega(H, \vartheta_p) - (1-\gamma)\beta g(\vartheta_p) f^p(\vartheta_p)\Big] \frac{1}{\alpha} \int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{f^k(\vartheta_k|H, 1)}{u'(c_2^k(H, \vartheta_p, \vartheta_k))} d\vartheta_k = \frac{\omega(H, \vartheta_p)}{u'(c_1^p(H, \vartheta_p))} \int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{\varphi(H, \vartheta_p, \vartheta_k)}{\beta_2^1 \omega(H, \vartheta_p)} d\vartheta_k$$

From (19), $\int_{\underline{\vartheta}_k}^{\overline{\vartheta}_k} \frac{\varphi(H,\vartheta_p,\vartheta_k)}{\beta_2^1 \omega(H,\vartheta_p)} d\vartheta_k = 1$. From FOC (10), $\omega(H, \vartheta_p) = \frac{\lambda g(\vartheta_p) f^p(\vartheta_p)\beta}{u'(c_1^k(H,\vartheta_p))}$. Hence,

$$\frac{1}{u'(c_1^p(H, \vartheta_p))} = \left[1 - \frac{(1-\gamma)u'(c_1^k(H, \vartheta_p))}{\lambda}\right] \mathbb{E}_{\vartheta_k} \left[\frac{1}{\alpha u'(c_2^k(H, \vartheta_p, \vartheta_k))}\right]$$

Using proposition 1's result: $\mathbb{E}_{\vartheta_k} \frac{u'(c_1^k(H,\vartheta_p))}{\alpha u'(c_2^k(H,\vartheta_p,\vartheta_k))} = 1$, therefore:

$$\frac{1}{u'(c_1^p(H, \vartheta_p))} = \mathbb{E}_{\vartheta_k} \left[\frac{1}{\alpha u'(c_2^k(H, \vartheta_p, \vartheta_k))}\right] - \frac{1-\gamma}{\alpha\lambda}$$

Intergenerational Wedge $t = 1$**: Low Type**

Same steps as for High Type, using FOCs (5) , (7), (11), and (13).

Bibliography

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2018, July). Education policy and intergenerational transfers in equilibrium. Technical report, Institute for Fiscal Studies.

Anger, S. and G. Heineck (2010, June). Do smart parents raise smart children? The intergenerational transmission of cognitive abilities. *Journal of Population Economics 23*(3), 1105–1132.

Belley, P. and L. Lochner (2007). The changing role of family income and ability in determining educational achievement. *Journal of Human Capital 1*(1), 37–89.

Castex, G. (2017). College risk and return. *Review of Economic Dynamics 26*, 91–112.

Cunha, F. and J. J. Heckman (2007). Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling. *Labour Economics 14*(6), 870–893.

Dynarski, S. M. (2003, March). Does aid matter? measuring the effect of student aid on college attendance and completion. *American Economic Review 93*(1), 279–288.

Farhi, E. and I. Werning (2007). Inequality and social discounting. *Journal of Political Economy 115*(3), 365–402. Lead article.

Farhi, E. and I. Werning (2010). Progressive estate taxation. *Quarterly Journal of Economics 125*(2), 635–673.

Findeisen, S. and D. Sachs (2016). Education and optimal dynamic taxation: The role of income-contingent student loans. *Journal of Public Economics 138*(C), 1–21.

Gale, W. G. and J. K. Scholz (1994, December). Intergenerational transfers and the accumulation of wealth. *Journal of Economic Perspectives 8*(4), 145–160.

Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003). Optimal indirect and capital taxation. *The Review of Economic Studies 70*(3), 569–587.

Golosov, M., M. Troshkin, and A. Tsyvinski (2016, February). Redistribution and social insurance. *American Economic Review 106*(2), 359–86.

Grochulski, B. and T. Piskorski (2010, May). Risky human capital and deferred capital income taxation. *Journal of Economic Theory 145*(3), 908–943.

Kapicka, M. and J. Neira (2015). Optimal Taxation with Risky Human Capital. Discussion Papers 1504, University of Exeter, Department of Economics.

Lochner, L. J. and A. Monge-Naranjo (2014). Student loans and repayment: Theory, evidence and policy.

Mirrlees, J. A. (1971). An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies 38*(2).

Paluszynski, R. and P. Yu (2019, 01). Optimal taxation with risky human capital and retirement savings. *SSRN Electronic Journal*.

Pavoni, N. and H. Yazici (2016). Intergenerational Disagreement and Optimal Taxation of Parental Transfers. *The Review of Economic Studies 84*(3), 1264–1305.

Stantcheva, S. (2015a). Optimal income, education, and bequest taxes in an intergenerational model. Working Paper 21177, National Bureau of Economic Research.

Stantcheva, S. (2015b, May). Optimal Taxation and Human Capital Policies over the Life Cycle. Nber working papers, National Bureau of Economic Research, Inc.