



**University of  
Zurich** <sup>UZH</sup>

University of Zurich  
Department of Economics

Working Paper Series

ISSN 1664-7041 (print)  
ISSN 1664-705X (online)

---

Working Paper No. 40

# **Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans**

Sebastian Findeisen and Dominik Sachs

Revised version, September 2012

---

# Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans\*

**Sebastian Findeisen**  
University of Zurich

**Dominik Sachs**  
University of Konstanz

First version: October 21, 2011  
This version: September 20, 2012

## Abstract

We study Pareto optimal tax and education policies when human capital upon labor market entry is endogenous and individuals face wage uncertainty. Though optimal labor distortions are history-dependent, i.e. depend on income and education, simple policy instruments can yield the desired distortions: a single nonlinear labor income tax schedule combined with income-contingent loans. To take the model to the (US) data, we simplify the model to a binary education decision (graduating from college or not). We find that for low and intermediate incomes the labor supply decision of college graduates should be distorted more heavily than for individuals without a college degree. As a consequence, the optimal student loan repayment schedule increases in income for this range. This result holds along the Pareto frontier. We compare the second best to a situation where loan repayment is restricted to be independent from income and find significant welfare gains.

*JEL-classification:* H21, H23, I21

**Keywords:** Optimal dynamic taxation, Education, Implementation

---

\*Contact: [sebastian.findeisen@uzh.ch](mailto:sebastian.findeisen@uzh.ch), [dominik.sachs@uni-konstanz.de](mailto:dominik.sachs@uni-konstanz.de). We owe special thanks to our advisors Friedrich Breyer and Fabrizio Zilibotti for ongoing support and valuable comments. We also thank Manuel Amador, Dan Anderberg, Carlos da Costa, Emmanuel Farhi, Mike Golosov, Bas Jacobs, Normann Lorenz, Elena Mattana, Florian Scheuer, Dirk Schindler, Kjetil Storesletten, Aleh Tsyvinski, Iván Werning and Christoph Winter for encouragement and helpful discussions. We also like to thank audiences at IIPF (August 2012, Dresden), Munich (July 2012), NORMAC (August 2012, Stroemstadt) PET (June 2011, Bloomington), SED (June 2012, Limassol), SMYE (April 2012, Mannheim), Stanford (October 2011), Yale (March 2011) and VFS (September 2012, Goettingen) for their comments. We thank Stefan Voigt for valuable research assistance. Parts of this research were undertaken while Sebastian Findeisen was visiting Yale and Dominik Sachs was visiting Stanford. We are grateful for their hospitality. Sebastian Findeisen acknowledges the support of the University of Zurich (Forschungskredit of the University of Zurich, grant no. 53210603).

# 1 Introduction

The deepening wage and income inequality observed in many countries over the last several decades is often linked to changes in the relative returns to education. In a recent survey, Acemoglu and Autor (2011) note that the college premium has been rising steadily in the US for the last 30 years. The same holds true for many other advanced economies, as an increasing college wage premium has contributed to an increase in the dispersion of the distribution of wages and income (Krueger, Perri, Pistaferri and Violante 2010).

Motivated by this tight connection between economic inequality and differences in educational attainment, we tackle the optimal design of integrated income tax and education policy systems in this paper. We analyze a dynamic heterogeneous agents economy, which consists of an education and a working period. Unlike former papers in the literature, we assume *both*, that individuals are ex-ante heterogeneous *and* the educational investment is risky.<sup>1</sup> The resulting Pareto optimal labor distortions are *history-dependent*, implying that implicit tax rates should condition on education in addition to income. History-dependent allocations and labor wedges are a standard outcome in dynamic tax problems (Golosov, Tsyvinski and Werning 2007, Kocherlakota 2010). We show that in our setting Pareto optimal allocations can be implemented by simple policy instruments: income-contingent student loans in addition to standard taxes on labor earnings and savings.

We apply our model to the case of a binary education decision: going to college or entering the labor market directly after high-school. We use income data to calibrate our model. The main focus of this numerical exercise is on the following question: given that we know that income-contingent student loans implement the optimum, how should they vary with income? We find that the optimal repayment schemes are – along the Pareto frontier – increasing for low and intermediate income at an almost linear slope. A second question we address is: can we gauge the welfare gains of such policies relative to simpler policies? We compare the optimal policies to a scenario where repayment cannot be income-contingent and find significant welfare gains.

The idea of income-contingent student loans has been proposed as early as 1955

---

<sup>1</sup>With the exception of Bénabou (2002), no paper has looked at optimal policies in a framework with heterogeneity and uncertainty yet. His approach differs from ours in that he does not solve for the second-best but compares the use of progressive income taxes with progressive education subsidies.

by Milton Friedman (1955), who envisioned repayment amounts to be proportional to income, i.e. a linearly increasing repayment schedule. Several countries like the United Kingdom, Australia or New Zealand currently have similar systems in place, where repayment is proportional to income within a certain income range. In this sense, our model develops a second-best argument for the use of such income-contingent repayment rates. Importantly, to arrive at this conclusion, we do not impose any restrictions on policy instruments in the spirit of the dynamic approach to public finance. It is noteworthy that the implementation result itself is not necessarily a case for Friedman's proposal. In theory, repayment could also be decreasing in income if college students are already over-insured (as compared to the second-best optimal insurance) by the income tax; in such a case it would be optimal to lower their effective marginal tax rates by having a decreasing repayment schedule. Looking at real world income distributions, however, our Mirrleesian analysis yields the opposite normative conclusion: a robust numerical result is that for low and intermediate incomes, repayment should indeed be increasing in income.

In our analysis, we first concentrate on a general model, where the focus lies on theory. We study an environment characterized by private information, which evolves stochastically over time. Individuals are born with different innate abilities and decide on their human capital investment. After obtaining education, agents learn their skill level that can be interpreted as wage, they work and generate income. The planner cannot observe innate ability, work effort and the skill level (wage). However, we do assume that the planner can observe educational investment. Labor market risk is taken into account in our model by the fact that individuals face a distribution of wages before entering the labor market. This distribution function depends on acquired human capital and innate ability. When deriving analytical results, we try to impose minimal structure on this stochastic process and work with a general distribution function. These modelling assumptions are quite flexible and in line with empirical evidence as we argue in more detail when introducing the model.

As hinted at above, the optimal allocation features history-dependent labor wedges, which is standard in dynamic models with informational frictions. We derive a simple formula for the optimal labor wedge that can be decomposed into two terms: one that is equivalent to the seminal formulas of Mirrlees (1971) and Saez (2001) and one that captures the optimal provision of dynamic incentives. We show that the dynamic

incentive term has a very simple intuition, related to how responsive the distribution of skills is to innate abilities.

Our theory also presents novel results on how individual education decisions are distorted by government intervention in constrained Pareto optimal allocations. This depends on two forces. First, implicit education subsidies are used to offset the distortionary impact of labor taxes on educational decisions. Intuitively, the gains from investing into education are partly reaped by the government via income taxes. Therefore the government uses an education subsidy to internalize this fiscal externality. Second, the planner wants to tax education in case he wants to redistribute from high to low innate ability types at the margin. This tax effect is increasing in the complementarity between innate ability and education.

Next, we move on to suggesting ways how Pareto optima can be decentralized. We show that the following policy instruments can implement the desired allocations: student loans with income-contingent repayment, a labor income tax schedule that solely depends on income and a nonlinear savings tax. The labor wedge consists of the marginal tax rate on labor income and the marginal loan repayment. The education wedge is implemented by the change in loan size as a function of education and how the repayment schedule changes with the loan size. In addition, we provide an alternative implementation with education-dependent income taxes similar to the idea of a graduate tax which has been proposed in the popular debate in the United Kingdom.

Finally, we move on from theory and present a simple application of the model, featuring a binary education decision and only two different ex-ante types. Importantly, this enables us to use estimates of factual as well as counterfactual income distributions from the labor economics literature (Cunha and Heckman 2007, 2008), both of which are needed as an input to simulate optimal policies. We follow Saez (2001) and append Pareto tails to the distributions of income before we calibrate the skill distribution. Under the – we believe – reasonable assumption that the college income distribution has thicker tails, the repayment schedule is increasing in income for a Utilitarian planner. Importantly, this result is stronger than that, as it holds along the Pareto frontier. Under the more conservative assumption that top incomes are distributed according to the same Pareto parameter, the repayment schedule is still increasing for low to intermediate incomes and roughly flat afterwards.

We gauge the welfare gains relative to a situation where labor wedges are constrained to be a function of *income only* but the planner can otherwise fully optimize. Concerning the policy instruments mentioned above, this is equivalent to restricting loan repayment to be independent of income. The key simplifying assumption to make this comparison tractable is the absence of income effects. We show that in this case, constraining labor wedges to be a function of income only is equivalent to restricting income to be a function of the wage only. Comparing the second best to this solution, we find welfare gains ranging from about 0.2% to 0.6% of lifetime consumption and we show how these gains change with risk-aversion and societal preferences for redistribution. Our welfare comparison is also a methodological contribution because we are the first to use a first-order approach for the case where labor taxes are constrained to depend on current income only and can therefore use a very fine type space.

**Related Literature.** Several previous papers have studied problems of optimal income taxation and their relation to education decisions. One strand of inquiry has worked under the assumption of *ex-ante homogeneity* and *risky* human capital. Da Costa and Maestri (2007) show that human capital should always be encouraged in the second-best optimum. Anderberg (2009) emphasizes that the risk properties of human capital are crucial for the question whether and how education should be distorted relative to a first-best rule.<sup>2</sup>

In a static setting with *heterogeneity* but without *uncertainty*, Bovenberg and Jacobs (2005) analyze how endogenous education alters the result of the Mirrleesian tax problem and conclude that distortions on the education margin through income taxation should be offset by education subsidies. Relatedly, Bohacek and Kapicka (2008) study a dynamic model with certainty and obtain similar results.<sup>3</sup>

Our paper is connected to the New Dynamic Public Finance literature. Concerning the characterization of Pareto optimal *allocations*, it is related to two recent contribu-

---

<sup>2</sup>Focusing on linear policy instruments Anderberg and Andersson (2003) as well as Jacobs, Schindler and Yang (2012) also discuss the importance of the risk properties of human capital.

<sup>3</sup>Kapicka (2006) introduces non-observable endogenous human capital into a dynamic, non-stochastic Mirrlees model where taxes can only be conditioned on current income. He shows that marginal tax rates are lowered due to the education margin. Bovenberg and Jacobs (2005) consider observable and unobservable investment and provide an interesting discussion about different implications. Grochulski and Piskorski (2010) focus on the implications of unobservable human capital investment for capital taxation in an ex-ante homogeneous agent setting with uncertainty. Relatedly, Shourideh (2012) studies optimal taxation of capital income with heterogeneous agents and capital income shocks.

tions that have analyzed optimal labor wedges in dynamic Mirrlees economies without endogenous education but with productivity shocks in every period, which we abstract from. Farhi and Werning (2012) characterize the evolution of the labor wedge over time analytically and numerically in such an environment. Golosov, Troshkin and Tsyvinski (2011) derive and illustrate optimal labor wedges in the cross section, connecting their generalized formulas to the static Mirrlees-Saez expressions.

Concerning the implementation of history-dependent allocations, our paper is related to Golosov and Tsyvinski (2006) who consider an environment with absorbing disability shocks and present an implementation in which disability insurance conditions on asset testing.<sup>4</sup> Similar as Golosov and Tsyvinski (2006), we analyze a model with a relatively simple stochastic environment, which enables us to present decentralizations of Pareto optima with simple policy instruments. Further, our implementation builds on recent work of Werning (2011), who shows that decentralized implementations of incentive compatible allocations exist, which are characterized by non-linear capital taxes that do not condition on current income and can even be history-independent.

The dependency of the tax schedules on characteristics other than income is related to the idea of *tagging*, as proposed by Akerlof (1978).<sup>5</sup> Whereas our analysis shares the feature of this method in that the social planner uses additional information to tailor marginal tax rates for each education level to the respective skill distribution, it differs substantially in the sense that education is not an immutable *tag*, but rather an endogenous variable.<sup>6</sup>

Finally, the paper is related to Luttmer and Zeckhauser (2008), who consider a setting in which agents have some form of private information about a distribution of outcomes. They apply this idea numerically to the case of income risk and two ex-ante types who self-select into two different tax schedules based on a private signal about their ability. For this purpose the authors interpret the college entry decision as a proxy for this private signal. In contrast, in our model the distribution of labor

---

<sup>4</sup>Kocherlakota (2005) provides an implementation based on income-contingent wealth taxes in a very general stochastic environment. Albanesi and Sleet (2006) derive an implementation with wealth-contingent income taxes for the case of *iid* shocks.

<sup>5</sup>More recently tagging is investigated by Cremer, Gahvari and Lozachmeur (2010), Mankiw and Weinzierl (2010) as well as Weinzierl (2011b).

<sup>6</sup>A similar logic arises in the recent paper of Scheuer (2012) in a model with endogenous occupational choices.

market skills depends on innate ability and education jointly and importantly, the education decision of individuals is endogenous.

This paper is organized as follows. Section 2 contains the basics of the model. In Section 3, we investigate dynamic incentive compatibility and describe the major properties of constrained efficient allocations. Decentralized implementations of constrained efficient allocations are provided in Section 4. We apply our model to the case of a binary education decision in Section 5. Section 6 concludes.

## 2 The Model

### 2.1 Technology and Preferences

Individuals, whose mass is normalized to one, live for two periods. In the first period they acquire human capital and in the second period they work. Individual labor market ability in period two is stochastic and the distribution depends on educational investment and initial type. We now formalize these ideas.

Individuals differ in innate ability  $\theta$ , which can be interpreted as a one dimensional aggregate of (non-)cognitive skills, I.Q. and family background, and is distributed in the interval  $[\underline{\theta}, \bar{\theta}]$  according to the cumulative density function (cdf)  $F(\theta)$ . After individuals learn their type  $\theta$ , which is private information, they make an educational investment  $z$ . Flow utility in period 1 is assumed to be  $u(c_1)$ , so we abstract from any effort costs of education.

In period two individuals draw their labor market ability  $a$  from a continuous conditional cdf  $G(a|z, \theta)$ , which depends on *innate* ability  $\theta$  and education  $z$  and has bounded support  $[\underline{a}, \bar{a}]$ . When individuals learn their type  $a$ , they make a labor-leisure decision. We assume that preferences in period 2 are given by:<sup>7</sup> $u(c_2) - \Psi\left(\frac{y}{a}\right)$ .

With this specification of the model, we capture many empirical regularities. First, assuming  $G(a|z, \theta)$  to be non-degenerate, our model captures the important fact of uncertainty in the labor market and risky educational investment. See e.g. Cunha and Heckman (2008) or Chen (2008) for recent contributions.

---

<sup>7</sup>We stick to separable preferences in the analytic part because the formula for the optimal labor wedge will be very easy to interpret. For the numerical simulations in Section 5, we sometimes depart from the separability assumption.



Second, we allow this cdf to be a function of innate ability  $\theta$  and thereby capture the fact that inequality in earnings is – to a certain extent – also determined by innate ability. Taber (2001) presents findings suggesting that much of the rise in the college premium may be attributed to a rise in the demand for unobserved skills, which are predetermined and independent of education. Indirect evidence for the importance of unobserved skills comes from the strong persistence of within education group inequality (Acemoglu and Autor 2011).

Third, the cdf  $G$  being a function of  $z$  captures the returns to education. Importantly, for most of our results, we do not impose a certain assumption on the pattern of these returns. Thus, our framework can either capture risk-increasing or risk-decreasing human capital investment.

Fourth, as long as  $\frac{\partial^2 G(a|z, \theta)}{\partial \theta \partial z} \neq 0$ , returns to educational investment differ in innate ability  $\theta$ . E.g., Carneiro and Heckman (2005) document that the returns can differ by as much as 19% points across individuals for one year of college. See also Lemieux (2006) for evidence on heterogeneity in returns.

To sharpen a few analytical results, it turns out helpful to place some structure on the behavior of  $G(a|z, \theta)$ :

**Assumption 1:**  $G(a|z', \theta) \succeq_{FOSD} G(a|z, \theta) \Leftrightarrow G(a|z', \theta) \leq G(a|z, \theta)$ , for all  $z < z'$  and for all  $(\theta, a)$ .

**Assumption 2:**  $G(a|z, \theta') \succeq_{FOSD} G(a|z, \theta) \Leftrightarrow G(a|z, \theta') \leq G(a|z, \theta)$ , for all  $\theta < \theta'$  and for all  $(z, a)$ .

**Assumption 3:**  $\frac{\partial^2 G(a|z, \theta)}{\partial \theta \partial z} \leq 0$  for all  $\theta, z$ .

These assumption will not be needed to derive our main results, but help to illustrate important aspects of the model. Whenever an assumption is needed for a result, we refer to it. The first and the second one capture the notion that education and innate ability should both have a direct effect on labor market outcomes represented by a first-order stochastic dominance shift; a rather natural way of ordering distributions. The third one captures their interaction and respects the compelling evidence of complementarity between early ability and educational investment.

When simulating optimal policies, we take into account that the education period is shorter than the working period and then consider  $T_e$  identical education periods

and  $T_w$  identical working periods. Hence, we basically just multiply period one by  $T_e$  and period two by  $T_w$ . This also implies that we abstract from further shocks to idiosyncratic labor productivity. This simplifies and helps to focus the analysis on the education-taxation link. In the empirical literature, there is no ultimate consensus on the relative importance of heterogeneity before labor market entry for lifetime inequality, but different approaches have attributed a major role to it. In recent work, Huggett, Ventura and Yaron (2011) estimate a structural life-cycle model and find that differences realized at the age of 23 can account for more of the variation in lifetime outcomes than do shocks received over the working lifetime. A standard reference is Kean and Wolpin (1997), who attribute an astonishing 90% to heterogeneity realized before labor market entry, while Storesletten, Telmer and Yaron (2004) estimate a number of about 50%.

## 2.2 The Laissez Faire Equilibrium

To lay out the basic properties of the model, we start with the characterization of the laissez-faire equilibrium without government intervention. In the second period, after agents have learned their labor market skill  $a$ , they choose labor supply, taking savings or private debt as given. This gives rise to the indirect utility function:

$$v_2(a, s(\theta)) = \max_{y, c_2} u(c_2) - \Psi\left(\frac{y}{a}\right) \quad \text{s.t.} \quad c_2 = y + Rs(\theta).$$

Individuals' utility functions are well-behaved –  $u(\cdot)$  is assumed to be increasing, at least twice continuously differentiable and strictly concave, and  $\Psi(\cdot)$  to be increasing, at least twice continuously differentiable and strictly convex. The parameter  $a$  is an individual's labor market skill, meaning that individuals with a higher  $a$  need to provide less labor effort to earn a given income  $y$ .

In the first period, agents decide how much to invest in education, and make a consumption-saving decision. Agents have access to a risk-free one period bond market; we impose no short-sale or enforcement constraints and an exogenous gross return  $R$ . This defines the indirect utility function:

$$V(\theta) = \max_{s, z, c_1} u(c_1) + \beta \int_{\underline{a}}^{\bar{a}} v_2(a, s) g(a|z, \theta) da \quad \text{s.t.} \quad c_1 + z = -s,$$

where  $s$  are savings. As already stated in the last section, we model the conditional distribution of skills  $g(a|z, \theta)$  as being determined by an agent's education level  $z$  and her innate ability  $\theta$ . Moreover, we focus on educational investment as a direct monetary cost. This is consistent with the idea that tuition fees and other monetary expenses are the most important factors on the cost side driving educational decisions. It is also in line with a foregone earnings interpretation, where more education delays labor market entry.  $z$  can be a sum of both factors.

We now present the main properties of the equilibrium without government intervention:

**Proposition 2.1.** *The laissez-faire allocation has the following properties:*

(i) *The Euler Equation holds:*

$$u'(c_1(\theta)) = \beta R \int_{\underline{a}}^{\bar{a}} u'(c_2(\theta, a)) dG(a|z(\theta), \theta)$$

(ii) *Labor supply is undistorted:  $\Psi' \left( \frac{y(\theta, a)}{a} \right) \frac{1}{a} = u'(c_2(\theta))$ .*

(iii) *The marginal cost of education is equalized to marginal benefits:*

$$u'(c_1(\theta)) = \beta \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) \frac{\partial g(a|z(\theta), \theta)}{\partial z} da$$

(iv) *If Assumptions 1-3 hold, educational investment is increasing and savings are decreasing in innate ability, i.e.  $z'(\theta) > 0$  and  $s'(\theta) < 0$ .*

*Proof.* See Appendix A.1 □

Parts (i)-(iii) follow directly from the first-order conditions. They are unsurprising properties, stating that private marginal rates of substitution are equated to technical marginal rates of transformation on the labor, capital, and education market.

Part (iv) states that without government policies, education and savings are monotone in innate ability  $\theta$  if Assumptions 1-3 are fulfilled. The proof provides instructive intuition for this result. It is sufficient to show that the objective is supermodular in all choice variables and type  $\theta$  (see Milgrom and Shannon (1994)). Inserting the budget constraint gives the problem reduced to two choices  $s$  and  $z$ :  $\max_{s, z} U(s, z; \theta, a, \beta) =$

$u(-s - z) + \beta \int_{\underline{a}}^{\bar{a}} v_2(a, s)g(a|z, \theta)da$ . This objective is supermodular in credit taken  $-s$ , education  $z$  and type if and only if:

$$\frac{\partial^2 U(s, z; \theta, \cdot)}{\partial s \partial \theta} < 0 \quad (1)$$

$$\frac{\partial^2 U(s, z; \theta, \cdot)}{\partial s \partial z} < 0 \quad (2)$$

$$\frac{\partial^2 U(s, z; \theta, \cdot)}{\partial z \partial \theta} > 0. \quad (3)$$

In Appendix A.1 we show that all inequalities hold. Equations (1) and (2) imply that the return to savings is lower for higher  $\theta$  types and with higher education, since expected labor skills are also higher. Equation (3) holds, since innate abilities and education are complementary to each other. Taken together the direct effects of being of higher type on credit and education are being reinforced by the relationship between the endogenous variables.

So far, we have assumed no limits on the ability of agents to borrow against future labor income. Imposing an ad hoc constraint of the form  $s \geq \phi$ , where  $\phi$  is some negative number, leaves most of the results from Proposition 2.1 unaffected. Notably, constrained agents will not be able to smooth consumption intertemporally as much as desired. Still education levels will be increasing in type:

**Corollary 2.2.** *Suppose Assumptions 1,2 and 3 hold. If agents face borrowing constraints  $s \geq \phi$ , education is monotone in type  $\theta$ , i.e.  $z'(\theta) > 0$  in the laissez-faire equilibrium.*

*Proof.* Above some threshold type, agents reach the borrowing limit and set  $s$  equal to  $\phi$ . Of those agents higher types still face the greater returns to education because of the complementarity and therefore choose a higher level of  $z$ .  $\square$

The empirical literature has documented sorting into education, based on heterogeneous expected returns (Cunha and Heckman 2007). The monotonicity of education in the laissez-faire equilibrium is consistent with that fact.

## 2.3 Wedges

For later purposes when we analyze optimal allocations and the respective tax and education finance systems that can implement such allocations, it is useful to define

three wedges. They are equal to implicit marginal tax rates on savings, labor income and education, respectively:

**Savings wedge:**

$$\tau_s(\theta) = 1 - \frac{u'(c_1(\theta))}{\beta R \int_{\underline{a}}^{\bar{a}} u'(c_2(\theta, a))g(a|z, \theta)da}$$

where  $R$  is the gross return on savings.  $\tau_s(\theta) > (<)0$  implies a downward (upward) distortion of savings.

**Labor wedge:** The labor wedge is positive (negative) if an individual works less (more) than it would at the intervention-free market price (which is her productivity level  $a$ ). Formally the labor wedge reads as:

$$\tau_y(\theta, a) = 1 - \frac{\Psi' \left( \frac{y(\theta, a)}{a} \right) \frac{1}{a}}{u'(c_2(\theta, a))}.$$

**Education wedge:** Here a positive (negative) wedge corresponds to an upward (downward) distortion of the education decision. Formally the education wedge reads as

$$\tau_z(\theta) = 1 - \frac{\beta \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) \frac{\partial g(a|z(\theta), \theta)}{\partial z(\theta)} da}{u'(c_1(\theta))}.$$

### 3 Constrained Pareto Optimal Allocations

In this section, we characterize constrained Pareto efficient allocations, where 'constrained' refers to the government being unable to observe agents' type  $\theta$  in period one and  $a$  in period two. In Subsection 3.1, we show that the problem is tractable using a first-order approach. In addition, we provide necessary as well as sufficient conditions for this approach to be valid. In Subsection 3.2, we analyze optimality conditions and their consequences for optimal policies. In Subsection 3.3, we explore the model using numerical simulations.

#### 3.1 Incentive Compatibility

We cast the problem as a sequential, dynamic mechanism – agents report an initial type  $\theta$  in the first period and, after uncertainty has materialized, report their productivity  $a$  in the second period. The planner assigns initial consumption levels  $c_1(\theta)$  and

education levels  $z(\theta)$  to individuals with innate ability  $\theta$ . Moreover, with each report there comes a sequence of utility promises for the next period  $\{v_2(\theta, a)\}_{a \in [a, \bar{a}]}$ . In the second period, the screening takes place over consumption levels  $c_2(\theta, a)$  and labor supply  $y(\theta, a)$ . All these quantities define an allocation in the economy. Dynamic incentive compatibility is ensured backwards, so we start analyzing the problem from the second period.

### 3.1.1 Second Period Incentive Compatibility

By the revelation principle, we can restrict attention to direct mechanisms. Suppose that in the first period agents have made truthful reports  $r_\theta(\theta) = \theta$ , although this is not necessary and just simplifies the exposition.<sup>8</sup> Conditions for this to be true are given in the next subsection. Conditional on this report, the second period incentive constraint must be met for any history of types  $(\theta, a)$  and reporting strategy  $r_a(a)$ :

$$u(c_2(\theta, a)) - \Psi \left[ \frac{y(\theta, a)}{a} \right] \geq u(c_2(\theta, r_a(a))) - \Psi \left[ \frac{y(\theta, r_a(a))}{a} \right] \quad \forall a, r_a(a), \theta.$$

Define the associated indirect utility function of the agents as:

$$v_2(\theta, a) = \max_{r_a(a)} u(c_2(\theta, r_a(a))) - \Psi \left[ \frac{y(\theta, r_a(a))}{a} \right].$$

Like in a standard Mirrleesian problem preferences satisfy single-crossing for given first-period reports. For global incentive compatibility it is, hence, necessary and sufficient that all local envelope conditions hold:

$$\frac{\partial v_2(\theta, a)}{\partial a} = \Psi' \left( \frac{y(\theta, a)}{a} \right) \frac{y(\theta, a)}{a^2}, \quad (4)$$

and the usual monotonicity condition, stating that  $y(\theta, a)$  is non-decreasing in ability levels  $a$ , is satisfied:

$$\frac{\partial y(\theta, a)}{\partial a} \geq 0. \quad (5)$$

---

<sup>8</sup>The reason is that in the second period the utility is a function of  $a, r_a(a)$  and  $r_\theta(\theta)$  but not of  $\theta$ .

### 3.1.2 First Period Incentive Compatibility

In the first period, an agent takes into account the effect of her report about  $\theta$  on future utility. First period incentive compatibility is ensured if and only if the double continuum of weak inequalities holds:

$$U(\theta, \theta) = u(c_1(\theta)) + \beta \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) dG(a|z(\theta), \theta) \geq$$

$$u(c_1(r_\theta(\theta))) + \beta \int_{\underline{a}}^{\bar{a}} v_2(r_\theta(\theta), a) dG(a|z(r_\theta(\theta)), \theta) = U(\theta, r_\theta(\theta)), \forall \theta, r_\theta(\theta),$$

where  $U(\theta, r_\theta(\theta))$  is the expected utility of an individual of type  $\theta$  reporting  $r_\theta(\theta)$ . The associated value function is:

$$V(\theta) = \max_{r_\theta(\theta)} u(c_1(r_\theta(\theta))) + \beta \int_{\underline{a}}^{\bar{a}} v_2(r_\theta(\theta), a) dG(a|z(r_\theta(\theta)), \theta). \quad (6)$$

By using the FOC of (6) one can easily derive the following envelope condition

$$\frac{dV(\theta)}{d\theta} = \beta \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) \frac{\partial g(a|z(\theta), \theta)}{\partial \theta} da, \quad (7)$$

which can easily be incorporated into a Lagrangian. As often done in screening problems, our strategy for solving the second-best problem is to work with a relaxed problem with only restrictions (4) and (7) being imposed and then check ex-post whether incentive compatibility is fulfilled. In the numerical explorations in Section 3.3 we find that incentive compatibility is always satisfied and therefore the first-order approach is valid for the primitives we consider.<sup>9</sup>

Next, we present a set of sufficient conditions.

**Lemma 3.1.** *Suppose Assumptions 2 and 3 hold, conditions (4), (5), (7) are satisfied and we have:*

$$(i) \frac{\partial y(\theta, a)}{\partial \theta} > 0,$$

---

<sup>9</sup>Our results of this section on dynamic incentive compatibility are related to previous work in the optimal non-linear pricing literature by Courty and Li (2000). They study optimal pricing schemes of a monopolist facing consumers with stochastic tastes. In our case the distribution of types tomorrow is endogenous since education is a choice. In recent contributions, Kapicka (2011) as well as Pavan, Segal and Toikka (2011) investigate the robustness and validity of the Mirrleesian first-order approach in a large class of general dynamic environments.

$$(ii) \frac{\partial z(\theta)}{\partial \theta} > 0,$$

then the considered allocation is incentive compatible.

*Proof.* See Appendix A.2. □

This lemma implies that instead of directly ex-post verifying whether period one incentive compatibility is satisfied in an allocation, one can alternatively check these two simple monotonicity conditions; if they are fulfilled, then the allocation is incentive compatible. Whereas condition (ii) is always fulfilled in our numerical examples, condition (i) was often violated for very low  $a$ ; we will comment on the reasons in Section 3.3 when we present numerical illustrations of the model.

### 3.2 Properties of Constrained Pareto Optimal Allocations

The planner maximizes

$$\int_{\underline{\theta}}^{\bar{\theta}} u(c_1(\theta)) d\tilde{F}(\theta) + \beta \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) dG(a|z(\theta), \theta) d\tilde{F}(\theta)$$

subject to (4), (7) and the resource constraint:

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[ c_1(\theta) - z(\theta) + \frac{1}{R} \int_{\underline{a}}^{\bar{a}} (c_2(\theta, a) - y(\theta, a)) dG(a|z(\theta), \theta) \right] dF(\theta) = 0.$$

We let the planner assign Pareto weights  $\tilde{F}(\theta)$  to individuals, depending (solely) on their initial skill level. Any distribution of these weights, which we normalize to satisfy  $\int_{\underline{\theta}}^{\bar{\theta}} \tilde{f}(\theta) d\theta = 1$ , corresponds to one point on the Pareto frontier.  $\lambda_R$  denotes the multipliers on the resource constraint and  $\eta(\theta)$  the multiplier function of the first-period envelope conditions. In Appendix A.3 the Lagrangian and the first-order conditions of the problem are stated.<sup>10</sup> We now characterize the wedges of second-best Pareto optimal allocations.

---

<sup>10</sup>As with the localization of the incentive constraints, we verify numerically that our solution fulfilling the first-order conditions indeed attains the maximum.



### 3.2.1 Labor Distortions

The following proposition characterizes the optimal labor wedge.<sup>11</sup>

**Proposition 3.2.** *At any constrained Pareto optimum, labor wedges satisfy:*

$$\frac{\tau_y(\theta, a)}{1 - \tau_y(\theta, a)} = \frac{1 + \varepsilon^u(\theta, a)}{\varepsilon^c(\theta, a)} \frac{u'(c_2(\theta, a))}{ag(a|z(\theta), \theta)} [\mathcal{A}(\theta, a) + \mathcal{B}(\theta, a)],$$

where

$$\begin{aligned} \mathcal{A}(\theta, a) = & G(a|z(\theta), \theta) \left[ \int_a^{\bar{a}} \frac{1}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta) \right. \\ & \left. - \frac{1 - G(a|z(\theta), \theta)}{G(a|z(\theta), \theta)} \int_{\underline{a}}^a \frac{1}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta) \right] \end{aligned}$$

$$\mathcal{B}(\theta, a) = \frac{1}{f(\theta)\lambda_R} R\beta \frac{\partial [1 - G(a|z(\theta), \theta)]}{\partial \theta} \eta(\theta),$$

where  $\varepsilon^u(\theta, a)$  ( $\varepsilon^c(\theta, a)$ ) is the uncompensated (compensated) labor supply elasticity of type  $(\theta, a)$  and

$$\eta(\theta) = \tilde{F}(\theta) - \frac{\int_{\underline{\theta}}^{\theta} \frac{1}{u'(c_1(\theta))} f(\theta) d\theta}{\int_{\underline{\theta}}^{\theta} \frac{1}{u'(c_1(\theta))} f(\theta) d\theta}.$$

*Proof.* See Appendix A.4.2. □

Elasticities play a double role for the optimal labor wedge. On the one hand, a higher *compensated* elasticity increases the excess burden of labor distortions and is therefore inversely related to optimal labor wedges; on the other hand, a higher *uncompensated* elasticity translates into higher income inequality for a given skill distribution, making insurance more valuable and therefore tends to increase optimal labor wedges. Moreover, the weighted mass  $ag(a|z(\theta), \theta)$  of agents whose labor supply is distorted by the tax is negatively related to the marginal tax reflecting a deadweight loss argument. The term  $u'(c_2(\theta, a))$  can be interpreted as capturing income effects –

<sup>11</sup>In a recent paper Golosov, Troshkin and Tsyvinski (2011) provide formulas for dynamic optimal labor wedges with exogenous human capital, connecting them to empirical observables in the spirit of the contributions of Diamond (1998) and Saez (2001) for the static Mirrlees model.

for individuals with low consumption income effects are stronger and therefore the disincentive effect of marginal tax rates is weakened.

Conceptually, the labor wedge consists of two parts in our dynamic economy. The first one,  $\mathcal{A}(\theta, a)$ , is a variation of the optimal tax formula in the static Mirrlees case, with the difference that for each initial type  $\theta$  there is one separate function  $\mathcal{A}(\theta, \cdot)$ .  $\mathcal{A}(\theta, a)$  disappears if agents are risk neutral and therefore second period insurance is not a concern. With risk-aversion, however, optimal policies provide insurance against the labor market risk agents face. Education enters through its effect on the conditional distribution of skills. The term is equivalent to the tax formula from the standard static Mirrlees problem with utilitarian welfare weights and, as shown in Appendix A.4.2, it can be rewritten as in Saez (2001).

The second term  $\mathcal{B}(\theta, a)$  is novel and shows how labor tax rates are used to optimally supply dynamic incentives. In contrast to  $\mathcal{A}(\theta, a)$ , it is independent of risk-preferences, but vanishes with ex-ante homogeneous agents. Fixing  $a$ , the implicit tax rate is proportional to  $\frac{\partial[1-G(a|z(\theta),\theta)]}{\partial\theta}$ , which measures the change in the probability of becoming a higher type than  $a$ . Higher initial types have, education constant, a higher probability of reaching a skill level above  $a$ . For two neighboring  $\theta$  the planner adjusts the labor wedges of the lower type to deter a deviation in the first period. The increase in the implicit marginal tax rate increases average taxes for all skills  $a^* \geq a$  and makes mimicking unattractive for the higher type. The effect is bigger, the more important the effect of innate abilities on labor market outcomes. The term also increases in  $\eta(\theta)$  which captures the redistributive preferences of the planner at the margin. Note that this term can also be negative, especially if Pareto weights are in favor of high  $\theta$ -types. Finally, a no-distortion at the top and bottom result goes through since  $\mathcal{B}(\theta, \bar{a}) = \mathcal{B}(\theta, \underline{a}) = \mathcal{A}(\theta, \bar{a}) = \mathcal{A}(\theta, \underline{a}) = 0$ .

### 3.2.2 Education Distortions

The following proposition characterizes optimal education policies.

**Proposition 3.3.** *At any constrained Pareto optimum, the education wedge is given by:*

$$\begin{aligned} \tau_z(\theta) = & \frac{1}{R} \int_{\underline{a}}^{\bar{a}} (y(\theta, a) - c_2(\theta, a)) \frac{\partial g(a|z(\theta), \theta)}{\partial z(\theta)} da \\ & + \frac{\beta\eta(\theta)}{\lambda_R f'(\theta)} \int_{\underline{a}}^{\bar{a}} \frac{\partial v_2(\theta, a)}{\partial a} \frac{\partial^2 G(a|z(\theta), \theta)}{\partial z(\theta) \partial \theta} da. \end{aligned}$$

*Proof.* See Appendix A.4.3. □

The first term captures the expected marginal fiscal gain of an increase in education. One can show that it is always positive under Assumption 1 (FOSD shift of education) and positive labor wedges. Investing a dollar more into education increases the expected obligation of an agent. The first part of the education wedge exactly offsets this effect from the labor wedge. Bovenberg and Jacobs (2005) have discovered this effect for the static Mirrlees model, whereas we show this fiscal externality part of the wedge extends to the setting with uncertainty, holding in expectation.

We now turn to the second term. Under Assumption 3 the cross-derivative  $\frac{\partial^2 G(a|z(\theta), \theta)}{\partial z(\theta) \partial \theta}$  is negative and  $\frac{\partial v_2(\theta, a)}{\partial a}$  is positive everywhere by second-period incentive compatibility. Further, for redistributive preferences  $\eta(\theta)$  is typically positive. Then the second part of the education wedge is negative and acts as an implicit tax on education. By distorting education downward, the planner relaxes binding incentive constraints and can redistribute more effectively in line with her preferences. This is a consequence of the complementarity assumption, stating that agents endowed with higher innate skills gain more from education at the margin. The bundle of a lower type, hence, becomes less attractive from the perspective of an agent if education is downward distorted. Such an intuition is familiar from the standard static Mirrlees model concerning positive marginal income tax rates on the interior of the skill set. Relatedly, for this incentive term a zero at the top and at the bottom  $(\underline{\theta}, \bar{\theta})$  result holds.<sup>12</sup>

---

<sup>12</sup>Jacobs and Bovenberg (2011) discuss deviations from a first-best rule for the education subsidy for a general earnings function in the case without uncertainty. Our result is similar to their first result that a complementarity in education and ability leads to a tax on education. They also consider the degree of complementarity between labor supply and education which might call for an education subsidy in contrast. This second effect is not captured in our environment since the returns to labor supply – once uncertainty has materialized – are independent from the education choice.

### 3.2.3 Savings Distortions

It turns out that the presence of education, which endogenously affects the probability distribution of tomorrow's skills, does not change the prescription of a positive intertemporal wedge, stemming from the optimality of the Inverse Euler equation in dynamic Mirrleesian models.<sup>13</sup> Some manipulations of the first-order conditions yield the following proposition:

**Proposition 3.4.** *In any constrained Pareto optimum, the inverse Euler equation holds:*

$$\frac{1}{u'(c_1(\theta))} = \frac{1}{\beta R} \int_{\underline{a}}^{\bar{a}} \frac{1}{u'(c_2(\theta, a))} g(a|z(\theta), \theta) da = \frac{1}{\beta R} E_{a|\theta} \left[ \frac{1}{u'(c_2(\theta, a))} \right].$$

*Proof.* See Appendix A.4.1. □

Jensen's inequality then implies  $\beta E [u'(c_2(\theta, a))] > u'(c_1(\theta))$  – the optimal allocation dictates a wedge between the intertemporal rate of substitution and transformation; savings are discouraged.

### 3.2.4 Multiple Education and Working Periods

We have worked with a two-period life-cycle model so far. We now briefly show how our results for optimal allocations can easily be extended to the case of multiple education and working life periods. For our numerical simulations, this enables us to account for the fact that the time spent in (tertiary) education relative to the working life is small.

Let  $T_e$  be the number of years education takes place and  $T_w$  the number of years an individual works, such that  $T_w + T_e = T$ . During education, so for the first  $T_e$  years in the life-cycle, consumption is constant and given by  $c_e(\theta)$ , consumption during the work period,  $c_w(\theta, a)$  and income earned  $y_w(\theta, a)$  are also constant, as can be shown under our assumption of separable preferences.

However, whereas there is no intertemporal wedge within the years of education and within one's working life, the inverse Euler equation still governs the relationship between  $c_e(\theta)$  and  $c_w(\theta, a)$  generating the familiar savings distortion. Assuming

---

<sup>13</sup> Diamond and Mirrlees (1978) and Rogerson (1985) were the first to derive it. In an important paper reviving the interest in the result, Golosov, Kocherlakota and Tsyvinski (2003) generalized it to a large class of dynamic environments, most importantly allowing for arbitrary skill processes.

$\beta R = 1$ , the labor wedges are given by the analogous expressions as in Proposition 3.4. Also the education wedge remains unaffected, except that the expression is adjusted for the time-horizon since the marginal benefits and incentive costs of education depend on the length of the working life period:

$$\begin{aligned} \tau_z(\theta) = & \sum_{t=T_e+1}^T \beta^{t-1} \int_{\underline{a}}^{\bar{a}} (y_w(\theta, a) - c_w\theta, a) \frac{\partial g(a|z(\theta), \theta)}{\partial z(\theta)} da \\ & + \frac{\sum_{t=T_e+1}^T \beta^{t-1} \eta(\theta)}{\lambda_R f(\theta)} \int_{\underline{a}}^{\bar{a}} \frac{\partial v_2(\theta, a)}{\partial a} \frac{\partial^2 G(a|z(\theta), \theta)}{\partial z(\theta) \partial \theta} da, \end{aligned}$$

with  $\beta R = 1$ .

### 3.3 Numerical Illustration

In this section we numerically explore our model in an illustrative manner. We consider two skill distributions as our primitives  $G(a|z, \theta)$  that lead to very similar equilibrium wage distributions and educational expenses for actual given policies. In one of the cases, the distribution function is characterized by a strong complementarity between innate skills and education. In the other, there is less complementarity and the direct effects of education and innate skills dominate.

We solve for the Utilitarian optimum, so  $\tilde{f}(\theta) = f(\theta) \forall \theta$ . The utility function is:

$$U(c, l) = \frac{c^{1-\rho}}{1-\rho} - \frac{(y/a)^\sigma}{\sigma},$$

where we set  $\sigma = 3$ , implying a Frisch elasticity of 0.5 and the CRRA coefficient to  $\rho = 2$ . The yearly risk-free rate is assumed to be 3% ( $R=1.03$ ) and  $\beta = 1/R$ .

We assume that labor market abilities are distributed log-normally following common practice and impose the location parameter  $\mu$  to be a function of  $\theta$  and  $z$ . Concerning  $\theta$ , we assume a uniform distribution within  $[0.1, 1]$ .

**Case (a) - Strong Complementarity:** The functional form of the location parameter is:

$$\mu(\theta, z) = 1.7 + 1.5\theta^{0.5}z^{0.15}.$$

In this case, individuals are the same if they do not acquire any education at all. However, the more education they acquire the stronger are the differences in the

location parameters. This inequality in  $\mu$  will be reinforced by the fact that agents have incentives to self-select into different education levels because of heterogeneous returns.

**Case (b) - Strong Direct Effect:** In the second case we assume,

$$\mu(\theta, z) = 1.5 + \theta + 0.75z^{0.25}.$$

In this case, individuals are already very different from the outset, i.e. if nobody acquires any education. The difference in the location parameter then stays constant for an uniform increase in  $z$  across agents. Although  $\frac{\partial^2 \mu(\theta, z)}{\partial \theta \partial z} = 0$  in this case, Assumption 3 is fulfilled for the relevant range. However, innate skills and education are weaker complements as compared to Case (a).

The respective parameters for the two cases as well as the respective constant marginal costs of education were chosen such that with an approximation of the current tax and college subsidy system in the US, the model roughly replicates per-capita expenditures on college education and the centers of the interval of the location parameters of the log-normal distributions is equal to the empirical value of the wage distribution.<sup>14</sup>

Figure 1 illustrates optimal education wedges for the two cases. In both cases, the optimal allocation features positive implicit education subsidies around 40%, which are relatively flat across innate types. The main difference between the two cases lies in the incentive effect. When innate skills and education are complements, the planner finds it optimal to tax education relative to a first best in line with Proposition 3.3. In Case (a) this incentive effect becomes as large as 17% whereas in Case (b) it hovers around zero.

Figure 2 illustrates the optimal labor wedges from Proposition 3.2. Panel (a) displays the optimal labor wedge as a function of income.<sup>15</sup> Darker regions refer to

<sup>14</sup>Following Gallipoli, Meghir and Violante (2011) we set the labor income tax to a flat rate of 27% and a lump sum transfer of one sixth of labor income per capita. We introduce a yearly education subsidy of 24%. In both cases, for these given policy instruments, average college education expenditures per year are roughly 30% of yearly median income; a long run average for the US (Gallipoli, Meghir and Violante 2011). The realized values of  $\mu(\theta, z)$  are within the range [2.02, 3.34], centered around 2.76, the value of the lognormal fit for the US wage distribution found by Mankiw, Weinzierl and Yagan (2009); as them, we set the scale parameter equal to 0.565.

<sup>15</sup>To economize on space we only show the figures for Case (b) here. The graphs for Case (a) turn out to look nearly identical.

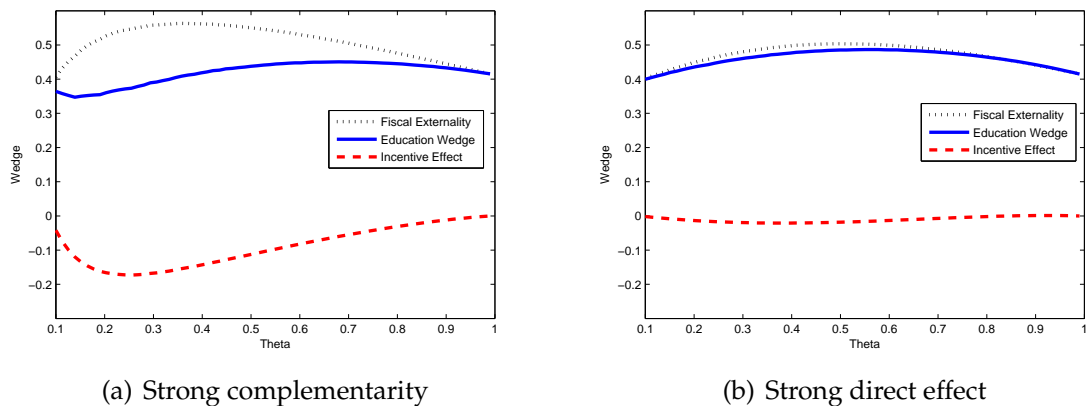


Figure 1: Optimal Education Wedges

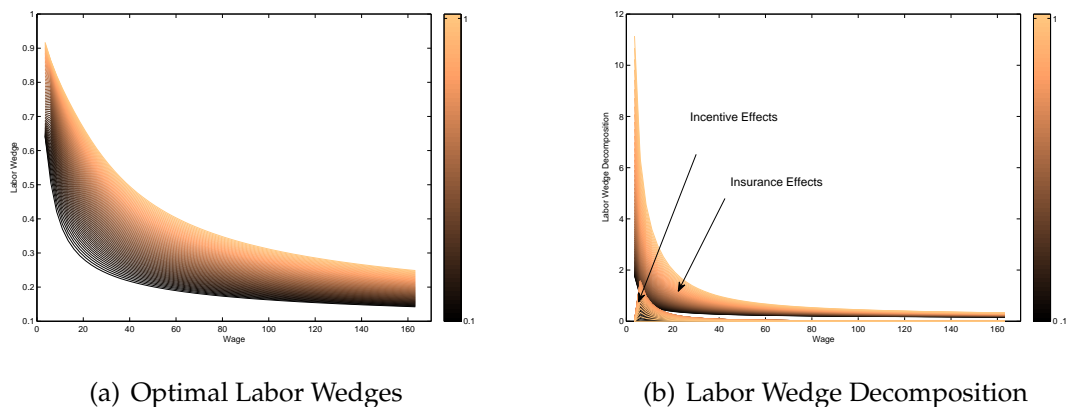


Figure 2: Optimal Labor Wedges

innate low types and lighter regions to innate high types. The picture shows that higher innate types face high labor wedges, whereas the shape of the wedges does not vary with  $\theta$ .<sup>16</sup> In the next panel (b), we illustrate the decomposition from Proposition 3.2 into the insurance term and the incentive term by plotting  $\mathcal{A}(\theta, a)$  and  $\mathcal{B}(\theta, a)$ . The set of insurance effects  $\mathcal{A}(\theta, a)$  lies above the set of incentive effects  $\mathcal{B}(\theta, a)$ . Still, especially at the beginning of the income distribution incentive effects contribute to higher implicit tax rates. The graphs also reveal that these incentive effects are of more importance for higher innate types on average.

<sup>16</sup>Since low incomes the distortions are strongly increasing in  $\theta$ , condition (i) of Lemma 3.1 is typically not fulfilled for low  $a$ .

## 4 Implementation

So far we only considered a direct mechanism, in which individuals make reports about their realized type and the planner assigns bundles of consumption, labor supply and education as functions of the reports. The focus in the characterization of the optimal allocation was on wedges or *implicit* price distortions of the allocation. In this section, we explore two decentralized implementations of constrained Pareto optima.

### 4.1 Implementation One: Student Grants and Income Taxes Conditioning on Education

#### 4.1.1 Two Periods

The benevolent government offers a menu of student grants to the agents. These grants  $\mathcal{G}$  are conditional on education. In the second period, there is a tax schedule in place, which, importantly, does not only condition on earnings but also on educational investment. Further, savings taxes are high enough to make private savings disappear from the market; the definition of the savings tax builds on Werning (2011). We summarize this in the following proposition:

**Proposition 4.1.** *Any constrained Pareto optimal allocation in a two period economy can be implemented by a grant schedule  $\mathcal{G}(z)$ , an education dependent income tax  $T(y, z)$  and a savings tax  $T^s(s)$ , where*

- $\mathcal{G}(z(\theta)) = z(\theta) + c_1(\theta)$
- $T(y(\theta, a), z(\theta)) = y(\theta, a) - c_2(\theta, a)$
- $T^s(s)$  as defined in Appendix A.5.

*Proof.* See Appendix A.5 □

**Implementation of savings wedges:** The savings function  $T^s(s)$  is prohibitively high such that all agents choose  $s = 0$ , hence in this implementation there are no private savings. However, as shown in Werning (2011) this comes without loss of generality: by a Ricardian equivalence argument, we can adjust  $\mathcal{G}(z(\theta))$  and  $T(y(\theta, a), z(\theta))$  with lump-sum transfers and deductibles to arrive with a non-linear savings tax



schedule, which produces non-zero private savings for every agent and the same allocation with the same distortion of consumption across periods. The full argument is found in Werning (2011).

**Implementation of labor wedges:** Agents enter the second period with no savings as argued above. Their budget constraint is then:  $T(y(\theta, a), z(\theta)) = y(\theta, a) - c_2(\theta, a)$ . From the agents' optimality conditions for  $y$  and  $c_2$  it follows that marginal tax rates  $T_y(y(\theta, a), z(\theta))$  are equal to labor wedges  $\tau_y(\theta, a)$  as characterized in Section 3.2.1.

**Implementation of education wedges:** To fix ideas, the budget constraints of an agent in both periods are given by:

$$\begin{aligned} c_1(\theta) + z(\theta) &\leq \mathcal{G}(z(\theta)) \\ c_2(\theta, a) &\leq y(\theta, a) - T(z(\theta), y(\theta, a)), \end{aligned}$$

where we already imposed the zero savings. In contrast to the optimal labor *wedge*, which equals the optimal labor *tax*, there is no single policy instrument for which the education wedge equals the marginal distortion of the policy. Instead, the government uses two instruments: i) the non-linear grant schedule  $\mathcal{G}(z)$ , which depends on education chosen and ii) the labor tax code in the second period. Using the agents' optimality conditions in the proposed implementation one can show that the wedge equals:

$$\tau_z(\theta) = \mathcal{G}'(z) - \int_{\underline{a}}^{\bar{a}} \frac{u'(c_2(\theta, a))}{u'(c_1(\theta))} g(a|z(\theta), \theta) T_z(y(\theta, a), z(\theta)) da$$

A positive value of  $\tau_z(\theta)$  encourages education at level  $\theta$ . The incentive for agents to increase their educational attainment comes from: i) An increase in their grant measured by  $\mathcal{G}'(z)$ <sup>17</sup> and ii) an increase or decrease in their labor income tax burden for all states, i.e.  $T_z(y(\theta, a), z(\theta))$ .

#### 4.1.2 T Periods

The possible extension to the life-cycle case is summarized as follows:

---

<sup>17</sup>Theoretically it could be the case that  $\mathcal{G}$  is (partly) decreasing in  $z$  if  $c_1(\theta)$  is sufficiently decreasing. However, this is rather unlikely and in all our numerical examples we have  $c_1'(\theta) > 0$ .

**Corollary 4.2.** *Any constrained Pareto optimal allocation in a  $T$ -period economy can be implemented by a grant schedule  $\mathcal{G}(z)$ , an education dependent income tax  $T_t(y^t, z)$  that conditions on the history of incomes and a savings tax  $T^s(s)$ .*

*Proof.* See Appendix A.6 □

The corollary is an application of insights from Werning (2007), who characterizes possible implementations in the dynamic deterministic Mirrlees problem. The problem here is similar to his model with the only difference that there is not only one tax schedule conditioning on the history of incomes  $y^t$ , but rather one for each  $\theta$ -type or each education level  $z$ .

One might wonder why the optimum is not simply implementable with history-independent schedules. The reason is that with such a tax system it might be possible for an individual to profit from tax arbitrage if tax schedules are sufficiently concave. Instead of earning the assigned income  $y$  every period, it can, e.g., then be favorable for an individual to earn  $y - \epsilon$  today and  $y + \epsilon$  tomorrow. Average gross income would be the same, however, due to concavity of  $T(y)$  the tax burden would decrease; if this effect is strong enough, it might compensate for the higher disutility of labor. In reality such strategic behavior to exploit the non-linearity of the tax system seems unlikely to occur. One reason is that shifting labor income between periods in such a way is often very costly or simply infeasible due to adjustment costs and hours restrictions.<sup>18</sup> Formally, let  $\tilde{C}(y^*)$  denote the present value of total adjustment costs of an individual that chooses income history  $y^*$ , and if  $y_{T_e+1}^* = y_{T_e+2}^* = \dots = y_T^*$ , then  $\tilde{C}(y^*) = 0$ , so if labor supply is constant, there are no adjustment costs.

**Definition** For a given adjustment cost function  $\tilde{C}(\cdot)$ , an income tax schedule  $T(y, z)$  is *tax arbitrage resistant* if  $V(y^*, c^*(y^*, T(y, z(\theta))), \tilde{C}(y^*)) \leq V(y^{truth}(\theta, a), c^{truth}(\theta, a)) \forall y^*$  and  $\forall (\theta, a)$ , where  $c^*(y^*, T(y), \tilde{C}(y^*))$  is the optimal consumption sequence given  $y^*$ ,  $T(y)$  and  $V(\cdot, \cdot)$  is the respective (deterministic) working life utility conditional on the realization of  $a$ .

With this in mind, we make the following assumption stating that strategic deviations are never feasible for a worker because of adjustment costs:

---

<sup>18</sup>For a recent treatment of hours constraint in the public economics literature see Chetty, Friedman, Olsen and Pistaferri (2011); for the implications of adjustment costs on hours choices and the labor supply elasticity see Chetty (2012). In addition to adjustment costs, there are also reasons outside our model, which make sophisticated intertemporal deviations unlikely, for example, idiosyncratic or aggregate uncertainty, imperfect credit markets or bounded rationality.

**Assumption 4:** The adjustment cost function  $\tilde{C}(\cdot)$  is such that the income tax schedules  $T(y, z)$  with  $T(y(\theta, a), z(\theta)) = y(\theta, a) - c_w(\theta, a)$  are tax arbitrage resistant.

This allows to implement constrained efficient allocations using a labor tax system that conditions on current income and one's education level only.<sup>19</sup> We summarize this in the following corollary:

**Corollary 4.3.** *Assume that Assumption 4 holds. Then any constrained Pareto optimal allocation in a  $T$ -period economy can be implemented by a grant schedule  $\mathcal{G}(z)$ , an education dependent income tax  $T(y, z)$  with  $T(y(\theta, a), z(\theta)) = y(\theta, a) - c_w(\theta, a)$  and a savings tax  $T^s(s)$ .*

*Proof.* By Assumption 4 we know that individuals prefer  $y^T(\theta, a)$  over any income sequence where  $y$  is not constant over time. By incentive compatibility we know that each individual prefers  $y^T(\theta, a)$  over  $y^T(\theta, a') \forall a', a$ ; combining this with the insights of Proposition 4.1 yields the result.  $\square$

## 4.2 Implementation Two: Student Loans with Income-Contingent Repayment

### 4.2.1 Two Periods

The previous implementation required that people with the same income but different levels of education pay different taxes. In reality people might perceive this as a violation of horizontal equity concerns, which could hinder the political feasibility of such policies. In this light we now present a more appealing alternative implementation with only one labor income tax schedule and a repayment scheme of the education grant.<sup>20</sup> Technically, this can be seen as a simple reinterpretation of the previous implementation – we take the tax system of the  $\underline{\theta}$  as the *common* labor income tax schedule and introduce an income-contingent repayment schedule, which conditions on the size of the loan.<sup>21</sup> Together both instruments are sufficient to replicate the optimal labor wedges. Formally we summarize this in the following proposition:

<sup>19</sup>We admit that this assumption might be considered as ad hoc. At the same time, the policy conclusions of *any* static nonlinear tax problem also share the problem of arbitrage opportunities once applied to (the dynamic) reality. Whereas arbitrage opportunities seem to be empirically important for capital income, we conjecture that making use of such arbitrage opportunities is of second-order importance for labor income.

<sup>20</sup>Diamond and Saez (2011) argue in a recent contribution that practical policy prescription from optimal tax models should not go against commonly held normative views (horizontal equity for example) and limit complexity to a reasonable degree. The second implementation seems in line with these recommendations.

<sup>21</sup>Related implementations are of course possible where the tax function of another  $\theta$ -type can be the labor income tax schedule in place. The extreme case would just be to say that income taxes do not exist and all

**Proposition 4.4.** *A constrained Pareto optimal allocation in a two-period economy can be implemented by a (compulsory) loan schedule  $L(z)$ , a loan repayment schedule  $\Gamma(y, L)$ , an income tax  $T(y)$  and a savings tax  $T^s(s)$ , where*

- $L(z(\theta)) = z(\theta) + c_1(\theta)$
- $\Gamma(y(\theta, a), L(z(\theta))) = c_2(\underline{\theta}, \tilde{a}(\underline{\theta}, y(\theta, a))) - c_2(\theta, a)$  if  $y \in [y(\underline{\theta}, \underline{a}), y(\underline{\theta}, \bar{a})]$  and  $\Gamma(y(\theta, a), L(z(\theta))) = y(\theta, a) - c_2(\theta, a)$  otherwise.
- $T(y) = y - c_2(\underline{\theta}, \tilde{a}(\underline{\theta}, y)) \forall y \in [y(\underline{\theta}, \underline{a}), y(\underline{\theta}, \bar{a})]$  and  $T = 0$  otherwise.
- $T^s(s)$  as defined in Appendix A.5.

where  $\tilde{a}(\theta, y)$  is the inverse of  $y(\theta, \cdot)$  for  $a$ .

*Proof.* The budget constraint of an individual reads as:

$$\begin{aligned} c_1(\theta) + z(\theta) &\leq L(z(\theta)) \\ c_2(\theta, a) &\leq y(\theta, a) - T(y(\theta, a)) - \Gamma(y(\theta, a), L(z(\theta))), \end{aligned}$$

which is equivalent to the budget constraint in Implementation 1 since by definition  $\mathcal{G}(z) = L(z) \forall z$  and  $T(y, z) = T(y) + \Gamma(y, z) \forall y, z$ . Hence it is a direct consequence of Proposition 4.1.  $\square$

The similarity to the other implementation is apparent. Optimal savings distortions are implemented as before. Using the agents' optimality conditions, one can show that the education wedge equals

$$\tau_z(\theta) = L'(z) - \int_{\underline{a}}^{\bar{a}} \frac{u'(c_2(\theta, a))}{u'(c_1(\theta))} g(a|z(\theta), \theta) \Gamma_L(y(\theta, a), L(z(\theta))) \frac{dL(z(\theta))}{dz} da,$$

and the labor wedge equals

$$\tau_y(\theta, a) = T'(y(\theta, a)) + \Gamma_y(y(\theta, a), L(z(\theta))).$$

Education wedges are implemented by the non-linear loans schedule and how repayment varies with education level in a risk-adjusted manner weighted by  $\frac{u'(c_2(\theta, a))}{u'(c_1(\theta))}$ .

---

schedules that were interpreted as history-dependent labor income schedules in implementation 1 can now be interpreted as repayment schedules. Taking the labor income tax schedule of the  $\underline{\theta}$ -type, however, seems to be more natural in our view. Especially in our application of the theory in Section 5.

The labor wedge is equal to the marginal tax rate and how loan repayment varies with income.

Finally, note that in Proposition 4.1, we imposed the loans to be mandatory. This is necessary because for some parameter specifications it might be possible that individuals might prefer not to take the government loan with income-contingent repayment but rather to borrow at the risk free rate  $R$ . This is rather likely if the income-contingent loan is such that on average one pays back much more than the loan size, i.e. the expected difference is more than the risk premium. To deal with this problem without making the income-contingent loans compulsory, one could add ex-ante participation constraints that guarantee voluntary participation in the loan program. Technically this would be similar to changing the Pareto weights in favor of those who would otherwise opt out of the loan program. As we will show in the application of the model in Section 5, our main qualitative result does not vary a lot along the Pareto frontier and we therefore conjecture that adding such participation constraints would not change that result significantly.

#### 4.2.2 T Periods

We will now show to what extent these results carry over to the case of  $T$  periods.

**Corollary 4.5.** *Any constrained Pareto optimal allocation in a  $T$ -period economy can be implemented by a (compulsory) loan schedule  $L(z)$ , a loan repayment schedule  $\Gamma_t(y^t, L)$ , an income tax  $T_t(y^t)$  and a savings tax  $T^s(s)$ .*

In the same way as Corollary 4.2 follows from Proposition 4.1, this Corollary follows from Proposition 4.4, we therefore omit a formal proof. Next, we use Assumption 4 to overcome the dependence of these policy instruments on the histories of incomes.

**Corollary 4.6.** *Assume that Assumption 4 holds. Then any constrained Pareto optimal allocation in a  $T$ -period economy can be implemented by a Loan schedule  $L(z)$ , a Loan Repayment Schedule  $\Gamma(y, L)$  that is constant over time, an income tax  $T(y)$  that is constant over time and a savings tax  $T^s(s)$ .*

Again we omit a formal proof as the result follows from Corollary 4.5 in the same way as Corollary 4.3 follows from Corollary 4.2. The implementation is appealing

since there is only one labor tax schedule and a yearly income-contingent repayment plan for each loan size.

## 5 An Application of the Model: College vs. High-School

We now present an empirically driven application of our model. We limit education to be a binary instead of a continuous choice. Agents either enter the labor market directly after high-school graduation or go to college before working. Additionally, we restrict the analysis to two different levels of innate ability.<sup>22</sup> These simplifications enable us to parameterize the model using factual and, importantly, estimated counterfactual earnings distributions from the empirical labor literature (Cunha and Heckman 2007, 2008). As laid out in Proposition 3.2, these counterfactual terms contribute to optimal marginal tax rates by deterring deviations in the first period. Further, the simplification has the advantage that it is easy to incorporate foregone earnings as an implicit cost of education.

### 5.1 Parametrization

Agents live for 49 years after they graduate from high-school (age 16-65). Afterwards they enter the labor market directly, or decide to go to college and graduate in four years. We label the two innate types  $\theta_{HS}$  and  $\theta_{CO}$  and assume that it is *a priori* optimal that  $\theta_{HS}$  chooses the lower educational attainment (high school) and that  $\theta_{CO}$  chooses the higher educational attainment (college). The incentive constraints read as:

$$\begin{aligned} & \sum_{t=0}^3 \beta^t u(c_E) + \sum_{t=4}^{48} \beta^t \int_{\underline{a}}^{\bar{a}} v_{CO}(a, \theta_{CO}) g(a|CO, \theta_{CO}) da \\ & \geq \sum_{t=0}^{48} \beta^t \int_{\underline{a}}^{\bar{a}} v_{HS}(a, \theta_{HS}) g(a|HS, \theta_{CO}), \end{aligned} \quad (8)$$

---

<sup>22</sup>This strong assumption implies that all individuals with the same education level are ex-ante equal, which is clearly not realistic. It can be seen as an indispensable approximation, which we need to make in order to identify innate types from real world data.

and

$$\begin{aligned} & \sum_{t=0}^{48} \beta^t \int_{\underline{a}}^{\bar{a}} v_{HS}(a, \theta_{HS}) g(a|HS, \theta_{HS}) da \\ & \geq \sum_{t=0}^3 \beta^t u(c_E) + \sum_{t=4}^{48} \beta^t \int_{\underline{a}}^{\bar{a}} v_{CO}(a, \theta_{CO}) g(a|CO, \theta_{HS}) da, \end{aligned} \quad (9)$$

where  $g(a|CO, \theta_{CO})$  and  $g(a|HS, \theta_{HS})$  are the probability density functions (pdfs) of the factual ability distributions and  $g(a|HS, \theta_{CO})$  and  $g(a|CO, \theta_{HS})$  are the pdfs of the counterfactual ability distributions. Note that college types now have to be compensated for their foregone labor earnings, the implicit cost of college education. To get the ability distributions, we take the factual and counterfactual earnings distributions for high-school graduates plotted in Cunha and Heckman (2007) in Figures 1 and 2.<sup>23</sup> After using a kernel smoother, we append Pareto tails at earnings of \$88,000. Finally, we smooth the resulting distribution again to overcome the kink from the appended tail. Given a (linear) approximation of the real world tax system we calibrate the implied skill distributions as input for the model from the optimality conditions of the agents.<sup>24</sup> We assume there is an atom of workers equal to five percent for each distribution reflecting unemployment or disability as in Mankiw, Weinzierl and Yagan (2009). The resulting calibrated skill distributions are illustrated in Figure 3. The share of high school and college types are set to 64.19% and 35.81%, respectively, as reported in Cunha and Heckman (2008). Following Gallipoli, Meghir and Violante (2011), we set the annual monetary cost of college education to roughly a third of median income in our data. We use the same flow utility function, the yearly discount factor and the interest rate as in Section 3.3.

---

<sup>23</sup>We used the software GetData Graph Digitizer to read out the data from the graphs. Since Cunha and Heckman (2007) consider the present value of lifetime earnings (18-65), we take a 47 years annuity with the same present value, i.e. we take something similar to average annual earnings.

<sup>24</sup>Saez (2001) has pioneered the approach to calibrate skills distributions from actual income distributions. We employ the same approximation as in the calibration of the full model in Section 3.3. This has been found to give a good fit to the US tax system by Jacquet, Lehmann and Van der Linden (2010).

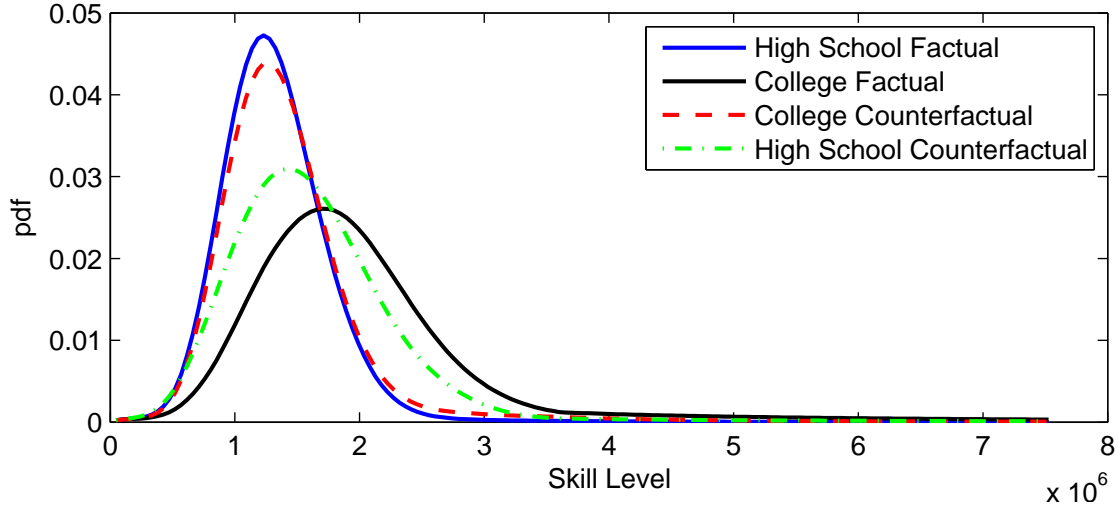


Figure 3: Skill Distributions

## 5.2 Simulations of Optimal Policies

### 5.2.1 Baseline: Thick College Tails and Utilitarian Welfare Function

To the best of our knowledge, there exists no systematic evidence on the conditional distributions of top incomes for college graduates and non-graduates separately. In the baseline case, we assume thicker tails for college graduates. For college graduates, we choose a Pareto parameter of 1.28, such that the ratio between the average income of individuals with income above the Pareto threshold is 4.5 relative to the threshold. For high-school graduates we use a Pareto parameter of 3, such that this ratio is 1.5. These values lie within the range of what has been typically found in empirical studies covering many countries and time periods (Atkinson, Piketty and Saez 2011). These top tails are not dependent on innate type  $\theta$  but are just determined by the education level. Importantly, we explore deviations from these assumptions on top incomes below. We start with the natural benchmark of a Utilitarian social welfare function but investigate optimal policies along the Pareto frontier below.

**Optimal Labor Wedges:** Figure 4(a) displays the optimal labor wedges as a function of yearly income up to \$300,000. Both schedules follow a U-shaped pattern, reflecting a seminal result from the static Mirrlees problem (Diamond 1998, Saez 2001). The intuition for the pattern is simple: for very low incomes, marginal distortions are high for two reasons: first distorting their labor supply is relatively harmless since



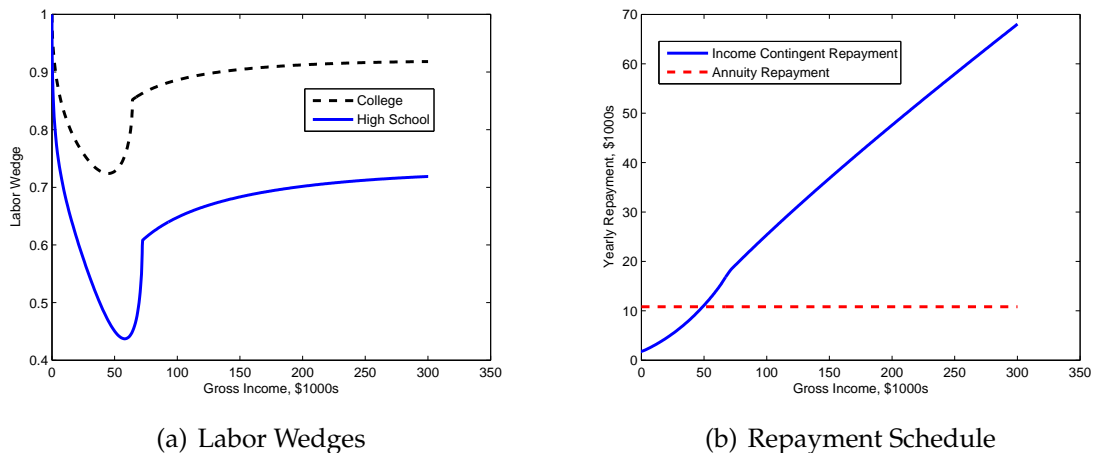


Figure 4: Utilitarian Optimum With Thick College Tails

they are rather unproductive. Second the inverse hazard rate  $\frac{1-G(a|\cdot,\cdot)}{g(a|\cdot,\cdot)}$  is rather high. Note that  $1 - G(a|\cdot,\cdot)$  is proportional to the additional revenue generated by the (implicit or explicit) marginal tax rate and  $g(a|\cdot,\cdot)$  is the mass of individuals whose labor supply is distorted. For intermediate incomes the density  $g(a|\cdot,\cdot)$  strongly increases making distortions more and more harmful, leading to a decrease in optimal distortions. Finally, due to the properties of the Pareto distribution, the ratio  $\frac{1-G(a|\cdot,\cdot)}{ag(a|\cdot,\cdot)}$  converges to a constant and as a consequence the labor wedges start to converge.<sup>25</sup>

Looking at Figure 3, one can see in which way tax distortions are tailored to the different skill/earnings distributions. At every point of the skill support before the Pareto tail kicks in, college labor distortions generate much bigger mechanical revenue effects for the government, which are reinforced by income effects. In the top income tails, the difference in the wedges also converges to a constant.<sup>26</sup>

**Repayment Schedule:** We now build on the implementation results from the previous section and illustrate optimal income-contingent repayment schedules. The (common) labor income tax schedule is determined by the high-school labor wedges. Figure 4(b) shows the yearly repayment of college debt as a function of income. The dotted line represents the annuity value of the loan, for comparison.

<sup>25</sup>See Saez (2001) for an elaborated discussion.

<sup>26</sup>All these results are related to the simulations of Luttmer and Zeckhauser (2008) who consider a static setting where going to college is purely a signal and not an investment; hence counterfactual and factual distributions are equal. One consequence of this is that labor wedges are very high for people without college degree but high income. The reason is that the effectiveness of deterring  $\theta_{CO}$ -types to mimic  $\theta_{HS}$ -types by increasing labor wedges of  $\theta_{HS}$ -types is increasing in the difference between the counterfactual distribution of the college type and the factual distribution of the high school type.

The slope of the repayment schedule is given by the difference in the labor wedges, as we laid out in the previous section. As the college wedge lies above the high-school wedge, repayment is increasing in income. In the tails the difference is constant and the repayment schedule becomes linear. Interestingly, also for lower income levels the slope of the schedule can be well approximated by a linear function.

As discussed at the end of Section 4.2.1, adding an ex-ante participation constraint to the problem that assures that the college type prefers the income-contingent loan over a loan from the private market should have a very similar effect like increasing the Pareto weight on the college type. Consequences for optimal policies should hence be similar as in Section 5.2.3 where we vary the Pareto weights

Relatedly, and as obvious from the Figure 4(b), graduates with income above US-\$60,000 pay back more than the (present) value of their loans in this example. It would also be interesting to introduce ex-post participation constraints, such that for no college graduate actual repayment exceeds the loan value by a certain amount. As repayment is strictly increasing in income, at least in our benchmark case, there would be a threshold, such that the constraint would be binding for all incomes above a certain level.

### 5.2.2 The Role of Pareto Tails

We now test if and how different assumption on top incomes across income distributions changes the results. First, we look at the case where all distributions (factual and counterfactual) share a common Pareto parameter of two. Figures 5(a) and 5(b) display the corresponding schedules. It shows that the college labor wedge still lies above the high-school labor wedge, for incomes up to roughly US-\$60,000, which leads to an increasing repayment schedule. Although the distribution share a common Pareto parameter, top labor wedges do not convergence to the exact same level because of the incentive effect we identified in the theory part and income effects. Top labor wedges become very similar, however, leading to repayment above US-\$60,000 being slightly U-shaped, but relatively flat.

Next, we use a Pareto parameter of 1.28 for the  $\theta_{CO}$ -type and 3 for  $\theta_{HS}$ -types – but in both cases independently from the college entry decision, so that the tails are *type-dependent*. The corresponding graphs (Figure 7(a) and 7(b)) appear in the appendix.

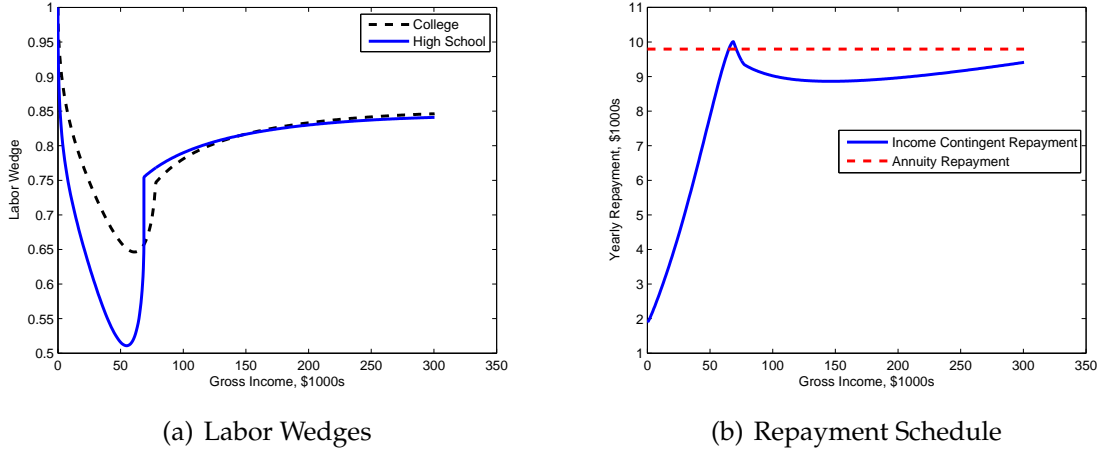


Figure 5: Utilitarian Optimum With Identical Tails

As is obvious from the figures, the schedules are almost identical to the *education-dependent* baseline case we studied in Section 5.2.1. This is driven by the fact that in both scenarios the equilibrium implies that the distribution for college graduates always has a thicker tail than the one for high-school graduates. What changes across the two scenarios are the counterfactual distributions, which only have a minor impact on the shape of the optimal schedules, however.

Lastly, we do not append a Pareto tail but work directly with the calibrated skill distributions. The graphs (Figure 8(a) and 8(b)) in the Appendix reveal that college labor wedges lie above their high school counterpart through the whole range of incomes, leading to an increasing repayment schedule.

### 5.2.3 Exploring the Pareto Frontier

We next aim to study optimal policies along the Pareto frontier by varying the weights  $\tilde{f}_{HS}$  and  $\tilde{f}_C$ , going back to our baseline specification with thick college tails. We start with the a case of a Rawlsian Planner, i.e.  $\tilde{f}_{HS} = 1$  and  $\tilde{f}_C = 0$ . As shown in the Appendix (Figures 9(a) and 9(b)), the general shape of the schedules is unaffected. This also holds for less redistributive ( $f_{HS} \leq \tilde{f}_{HS} \leq 1$ ) weights between the Rawlsian and Utilitarian points. We simulate cases where the planner puts a higher weight on college types, such that the incentive constraint binds in the other direction. Figures 10(a) and 10(b) in the Appendix show such a case, where all the weight is on the college type. Even in this extreme case the main conclusions remain unaffected.

### 5.3 The Welfare Gains From Income-Contingent Repayment

In the previous section, we have illustrated how optimal income-contingent repayment schedules may look like. As an important next step, we aim at quantifying what the potential welfare effects of income-contingency might be. The natural policy comparison is the case where repayment is not contingent on income and there is a single income tax schedule; i.e. concerning allocations, the restriction is that individuals with the *same income* should face the *same labor wedge*. Concerning education subsidies or savings taxes, we do not impose any restriction. We view this as the natural policy comparison since this is the situation that is true for most countries. The challenge is to incorporate this restriction – that the labor wedge should be a function of income only – into a Pareto problem. With income effects this problem is generally not tractable. Whenever individuals differ in their savings, individuals with the same wage will choose a different labor supply  $y$ ; hence it is not clear how to screen individuals given this restriction on policy instruments. In contrast, assuming away income effects by considering the following utility function

$$U(c, y, a) = u\left(c - \Psi\left(\frac{y}{a}\right)\right) \quad (10)$$

makes the problem tractable: in any implementation where labor wedges are only a function of current skills, we know that everyone with the same current skill  $a$  will choose the same level of income. Hence the restriction that the labor wedge is only a function of current income is *equivalent* to:<sup>27</sup>

$$y(\theta, a) = y(a). \quad (11)$$

Such a restriction can easily be incorporated into a Pareto problem. The following proposition states how a Pareto optimal allocation subject to (11) can be implemented in the binary model.<sup>28</sup>

---

<sup>27</sup>For the numerical simulations in this subsection, we assume that the allocation is constant within the working period. In general this might not be true for non-separable preferences. Since the main focus of our analysis is to compare history-independent with history-dependent instruments and this constant allocation assumption is made for both cases, we conjecture that it does not bias our result significantly.

<sup>28</sup>Naturally, another implementation of this optimum would involve a single labor tax schedule with differential lump-sums and education grants offered by the government.

**Proposition 5.1.** *Assume that preferences satisfy (10) and that the allocation is constant within the working period. Then any Pareto optimal allocation subject to private information and (11) can be implemented by a loan for college students  $L$ , a yearly loan repayment  $\Gamma$  and an income tax  $T(y)$  that is constant over time, where these policy instruments satisfy*

- $T(y(a)) = y(a) - c(\theta_{HS}, a)$
- $\Gamma = c(\theta_{HS}, a) - c(\theta_{CO}, a)$
- $L = \sum_{t=0}^3 \beta^t c_E + z.$

Note that there is no need for a savings tax since the intertemporal wedge remains undistorted between the college and the working period. Due to the absence of income effects and in contrast to the case with separable preferences, the standard inverse Euler equation argument for a savings distortion does not apply, although agents are risk-averse. The argument for a savings distortion due to nonseparable preferences as in Golosov, Troshkin and Tsyvinski (2011) also does not apply since in the education period utility is only given by consumption so there is no interaction with labor effort.<sup>29</sup> Finally, note that Assumption 4 is not needed since the tax arbitrage idea does not apply in the absence of income effects.

This theoretical result is very helpful for gauging the welfare gains, which stem from the income-contingency of the educational loans. If we compared our results with the case of linear policy instruments, for example, it would be unclear whether the welfare losses were mainly due to the inability of using history-dependent labor distortions (i.e. income-contingency) or from the inability of using nonlinear policy instruments. Weinzierl (2011a) follows a similar approach when estimating the welfare gains of age-dependent taxation comparing non-linear labor taxes across scenarios. Due to the quasi-linearity assumption on preferences, we can make use of a first-order approach to the incentive constraints and therefore look at a very fine type space exploiting more empirical knowledge about the skill distributions. Our welfare gains are also related to Farhi and Werning (2012) and Golosov, Troshkin and Tsyvinski (2011), who compare their optimal policies in a life-cycle framework with

---

<sup>29</sup>With an effort cost of education, which is not separable from consumption the savings wedge would be non-zero in general.

	Thick College Tails		Same Tails	
	Utilitarian	Rawlsian	Utilitarian	Rawlsian
$\rho \approx 1$	0.19%	0.26%	0.17%	0.16%
$\rho = 2$	0.37%	0.32%	0.32%	0.29%
$\rho = 3$	0.50%	0.46%	0.44%	0.47%
$\rho = 4$	0.62%	0.58%	0.54%	0.51%

Table 1: Welfare Gains from Income-Contingency

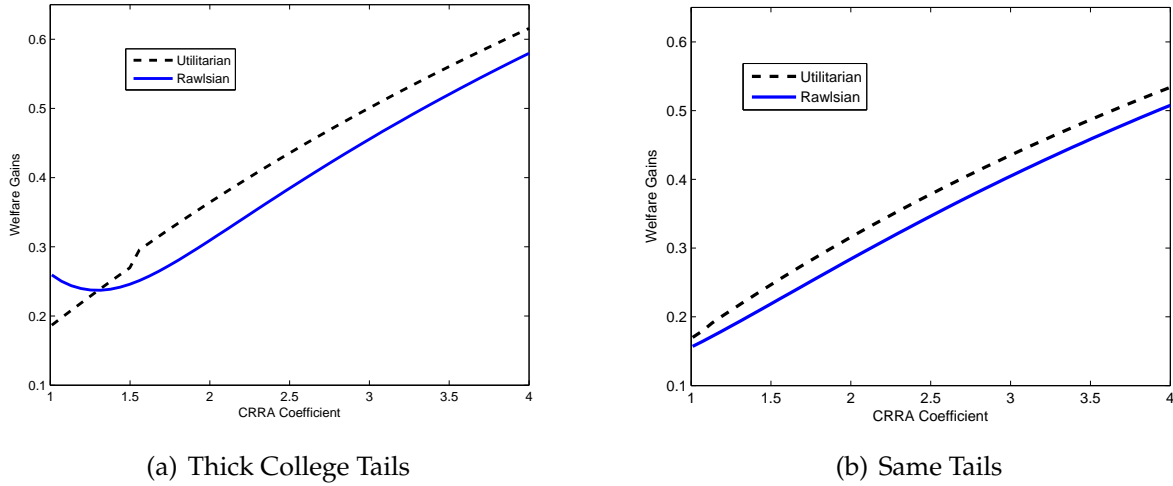


Figure 6: Welfare Gains as Function of Risk Aversion

age dependent and age independent linear taxes.

For the simulations we work with a CRRA specification so that:

$$U(c, y, a) = \frac{\left(c - \frac{(y/a)^\sigma}{\sigma}\right)^{1-\rho}}{1-\rho},$$

with  $\sigma = 3$  as before, implying a constant elasticity of 0.5. We aim to explore the welfare gains from income-contingent repayment schemes and how sensitive these gains are to different assumptions about risk-aversion and the social welfare function. In Table 1 and Figures 6(a) and 6(b) we present the consumption equivalent welfare gains as the CRRA parameter  $\rho$  varies from (slightly above) 1 to 4. For the

social welfare function, we focus on the two most popular cases in the literature, the Utilitarian case with  $f_{HS} = \tilde{f}_{HS}$  and  $f_C = \tilde{f}_C$  and the Rawlsian case with  $\tilde{f}_{HS} = 1$ . We assume thicker college tails as in the previous section in one case and – more conservatively – identical tails in the other case.

In the parametrization matching our baseline scenario with  $\rho = 2$ , thicker college tails and a Utilitarian planner, we find a respectable gain of 0.37% lifetime consumption.<sup>30</sup> Interestingly, the gains are almost everywhere higher for a Utilitarian planner. Intuitively, the individuals who gain the most from an income-contingent repayment scheme are college graduates with relatively low earnings. By assumption, the Rawlsian planner places no weight on college graduates, explaining why history dependence is actually more valuable for the Utilitarian planner. Next, the gains are increasing in risk-aversion almost across the whole support of  $\rho$ . This is to be expected for both the Utilitarian and the Rawlsian planner since history-dependent labor wedges in general improve insurance especially for low income states, which is valued more highly as risk-aversion increases.<sup>31</sup> Finally, comparing the left and right panel, as expected, the welfare gains from income-contingency are higher if the income distributions differ in their tails. With thicker college tails, it becomes more valuable for the planner to differentiate labor distortions. In line with this, the optimal *history-independent* tax schedules always lie in between the optimal college and high-school *history-dependent* labor wedges.

## 6 Conclusion

In this paper, we have studied the implications of endogenous education decisions before labor market entry on Pareto optimal tax policies in a dynamic environment with heterogeneous agents and uncertainty. An attractive way to decentralize Pareto optimal allocations is to have the government support students to finance consumption and tuition during education. During their working life students pay back these loans, contingent on income and loan size. Our paper therefore makes a second-best

---

<sup>30</sup>The numbers are almost identical when we take type-dependent tails.

<sup>31</sup>Over a small range with thick college tails for the Rawlsian planner the gains are slightly decreasing in risk-aversion. When risk-aversion is small, college types are more responsive to differential labor wedges for top incomes. Then it can be more important for the planner to tailor labor wedges according to the incentive effect described in Section 3. As our results show, however, this effect is second-order.

argument in favor of student loans with income-contingent repayment rates and, in addition, provides guidance for the optimal design of such repayment schedules.

We have put the model to work illustrating these policy recommendations for the case of a binary education choice: graduating from college or entering the labor market directly after high-school. We found the optimal repayment plan for college loans to be increasing at an almost linear rate for low and intermediate incomes, whereas the slope for top incomes hinges on the assumed distribution of top incomes. Interestingly, this qualitative result holds along a large part of the Pareto frontier. The welfare gains from using income-contingent repayment rates (history-dependent labor wedges) relative to non-contingent loans (labor wedges that are only a function of income) range from 0.2% to 0.6% of lifetime consumption for CRRA parameters ranging from one to four. Two key parameters for more definite policy conclusions are the two Pareto coefficients measuring the thickness of the tails for top incomes. It would be very valuable to obtain empirical estimates for different education levels to inform optimal policies. In addition, one could distinguish between different college majors, as the risk properties are probably very different across majors.

We have abstracted from several aspects that can be tackled in future work. First, we have abstracted from initial wealth heterogeneity. In an environment where individuals differ concerning the income and wealth of their parents, typically the question arises to what extent optimal education policies should depend on parents' income and wealth. Very related to this question, Gelber and Weinzierl (2012) have recently taken up the task of showing how the optimal history-independent tax system changes, when children's abilities depend on parents' financial resources. Second, due to our assumption that all labor market risk is realized directly after labor market entry, some aspects concerning the optimal timing of repayment were naturally disregarded. Relatedly, we did not consider human capital accumulation after labor market entry like on-the-job training. Third, we assumed full commitment to all policies from the government side. Relaxing these assumptions might be a fruitful area for future research.



## References

- ACEMOGLU, D., AND D. AUTOR (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics Vol. 4*, ed. by O. Ashenfelter, and D. Card, pp. 1043–1171. Elsevier.
- AKERLOF, G. (1978): "The economics of 'tagging' as applied to the optimal income tax, welfare programs, and manpower planning," *The American Economic Review*, 68(1), 8–19.
- ALBANESI, S., AND C. SLEET (2006): "Dynamic optimal taxation with private information," *The Review of Economic Studies*, 73(1), 1.
- ANDERBERG, D. (2009): "Optimal Policy and the Risk Properties of Human Capital Reconsidered," *Journal of Public Economics*, 93(9-10), 1017–1026.
- ANDERBERG, D., AND F. ANDERSSON (2003): "Investments in Human Capital, Wage Uncertainty, and Public Policy," *Journal of Public Economics*, 87(7-8), 1521 – 1537.
- BÉNABOU, R. (2002): "Tax and Education Policy in a Heterogenous-Agent Economy: What of Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 70(2), 481–517.
- BOHACEK, R., AND M. KAPICKA (2008): "Optimal Human Capital Policies," *Journal of Monetary Economics*, 55(1), 1–16.
- BOVENBERG, L., AND B. JACOBS (2005): "Redistribution and Education Subsidies are Siamese Twins," *Journal of Public Economics*, 89(11-12), 2005–2035.
- CARNEIRO, P., AND J. HECKMAN (2005): "Human Capital Policies," in *Inequality in America: What Role for Human Capital Policies*, ed. by J. Heckman, and A. Krueger, pp. 77–240. The MIT Press.
- CHEN, S. (2008): "Estimating the Variance of Wages in the Presence of Selection and Unobserved Heterogeneity," *Review of Economics and Statistics*, 90(2), 275–289.
- CHETTY, R. (2012): "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 80(3), 969–1018.

- CHETTY, R., J. FRIEDMAN, T. OLSEN, AND L. PISTAFERRI (2011): "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *Quarterly Journal of Economics*, 126(2), 749.
- COURTY, P., AND H. LI (2000): "Sequential Screening," *The Review of Economic Studies*, 67(4), 697–717.
- CREMER, H., F. GAHVARI, AND J. LOZACHMEUR (2010): "Tagging and Income Taxation: Theory and an Application," *American Economic Journal: Economic Policy*, 2(1), 31–50.
- CUNHA, F., AND J. HECKMAN (2007): "Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling," *Labour Economics*, 14(6), 870–893.
- (2008): "The Evolution of Labor Earnings Risk in the US Economy," *Working Paper, University of Chicago*.
- CUNHA, F., J. HECKMAN, AND S. NAVARRO (2005): "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economic Papers*, 57(2), 191–261.
- DA COSTA, C. E., AND L. J. MAESTRI (2007): "The risk Properties of Human Capital and the Design of Government Policies," *European Economic Review*, 51(3), 695–713.
- DIAMOND, P. (1998): "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 88(1), 83–95.
- DIAMOND, P. A., AND J. A. MIRRLEES (1978): "A model of Social Insurance with Variable Retirement," *Journal of Public Economics*, 10(3), 295–336.
- DIAMOND, P. A., AND E. SAEZ (2011): "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, 25(4), 165–90.
- FARHI, E., AND I. WERNING (2012): "Insurance and Taxation over the Life Cycle," *Review of Economic Studies* (forthcoming).
- FRIEDMAN, M. (1955): "The role of government in education," *Rutgers University Press*.

- GALLIPOLI, G., C. MEGHIR, AND G. VIOLANTE (2011): "Equilibrium effects of education policies: A quantitative evaluation," *Working Paper*.
- GELBER, A., AND M. WEINZIERL (2012): "Equalizing Outcomes vs. Equalizing Opportunities: Optimal Taxation when Children's Abilities Depend on Parents' Resources," *NBER Working Paper 18332*.
- GOLOSOV, M., N. KOCHERLAKOTA, AND A. TSYVINSKI (2003): "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 70(3), 569–587.
- GOLOSOV, M., M. TROSHKIN, AND A. TSYVINSKI (2011): "Optimal Dynamic Taxes," *NBER Working Paper No. 17642*.
- GOLOSOV, M., AND A. TSYVINSKI (2006): "Designing optimal disability insurance: A case for asset testing," *Journal of Political Economy*, 114(2), 257–279.
- GOLOSOV, M., A. TSYVINSKI, AND I. WERNING (2006): "New Dynamic Public Finance: A User's Guide," *NBER Macroeconomics Annual*, 21, 317–387.
- GROCHULSKI, B., AND T. PISKORSKI (2010): "Risky Human Capital and Deferred Capital Income Taxation," *Journal of Economic Theory*, 145(3), 908–943.
- HUGGETT, M., G. VENTURA, AND A. YARON (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101.
- JACOBS, B., AND A. L. BOVENBERG (2011): "Optimal Taxation of Human Capital and the Earnings Function," *Journal of Public Economic Theory*, 13(6), 957–971.
- JACOBS, B., D. SCHINDLER, AND H. YANG (2012): "Optimal Taxation of Risky Human Capital," *Scandinavian Journal of Economics*, 114(3), 908–931.
- JACQUET, L., E. LEHMANN, AND B. VAN DER LINDEN (2010): "Optimal Redistributive Taxation with both Extensive and Intensive Responses," *CESifo Working Paper 3308*.
- KAPICKA, M. (2006): "Optimal Income Taxation with Human Capital Accumulation and Limited Record Keeping," *Review of Economic Dynamics*, 9(4), 612–639.
- (2011): "Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-Order Approach," *Working Paper*.

- KEANE, M., AND K. WOLPIN (1997): "The career decisions of young men," *Journal of Political Economy*, 105(3), 473–522.
- KOCHERLAKOTA, N. (2010): *The new dynamic public finance*. Princeton Univ Press.
- KOCHERLAKOTA, N. R. (2005): "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 73(5), 1587–1621.
- KRUEGER, D., F. PERRI, L. PISTAFERRI, AND G. L. VIOLANTE (2010): "Cross Sectional Facts for Macroeconomists," *Review of Economic Dynamics*, 13(1), 1–14.
- LEMIEUX, T. (2006): "Postsecondary Education and Increasing Wage Inequality," *American Economic Review*, 96(2), 195–199.
- MANKIW, N., AND M. WEINZIERL (2010): "The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution," *American Economic Journal: Economic Policy*, 2(1), 155–176.
- MANKIW, N., M. WEINZIERL, AND D. YAGAN (2009): "Optimal Taxation in Theory and Practice," *Journal of Economic Perspectives*, 23(4), 147–174.
- MILGROM, P., AND C. SHANNON (1994): "Monotone Comparative Statics," *Econometrica*, 62(1), 157–180.
- MIRRLEES, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *The Review of Economic Studies*, 38(2), 175–208.
- PAVAN, A., I. SEGAL, AND J. TOIKKA (2011): "Dynamic Mechanism Design: Incentive Compatibility, Profit Maximization, and Informational Disclosure," *Working Paper*.
- ROGERSON, W. P. (1985): "Repeated Moral Hazard," *Econometrica*, 53(1), 69–76.
- SAEZ, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68(1), 205–229.
- SALANIÉ, B. (2003): *The Economics of Taxation*. The MIT Press.
- SCHEUER, F. (2012): "Entrepreneurial Taxation with Endogenous Entry," *Working Paper, Stanford University*.

- SHOURIDEH, A. (2012): “Optimal Taxation of Capital Income: A Mirrleesian Approach to Capital Accumulation,” *Working Paper, Wharton School*.
- STORESLETTEN, K., C. TELMER, AND A. YARON (2004): “Consumption and risk sharing over the life cycle,” *Journal of Monetary Economics*, 51(3), 609–633.
- TABER, C. (2001): “The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?,” *Review of Economic Studies*, 68(3), 665–691.
- WEINZIERL, M. (2011a): “The Surprising Power of Age-Dependent Taxes,” *The Review of Economic Studies*, 78(4), 1490–1518.
- (2011b): “Why Do We Redistribute So Much But Tag So Little? Normative Diversity and Optimal Taxation,” *Working Paper, Harvard Business School*.
- WERNING, I. (2011): “Nonlinear Capital Taxation,” *Working Paper, MIT*.

## A Appendix

### A.1 Properties of the Laissez-Faire Allocation

To show part (iv), we show that:

$$\frac{\partial^2 U(s, z; \theta, \cdot)}{\partial s \partial \theta} = -\beta \int_{\underline{a}}^{\bar{a}} \frac{\partial^2 v_2(a, s)}{\partial s \partial a} \frac{\partial G(a|z, \theta)}{\partial \theta} da \quad (12)$$

$$= -\beta \int_{\underline{a}}^{\bar{a}} \frac{\partial y(s, a)}{\partial s} \left[ \Psi' \left( \frac{y(s, a)}{a} \right) \frac{1}{a^2} + \Psi'' \left( \frac{y(s, a)}{a} \right) \frac{y(s, a)}{a^2} \right] \frac{\partial G(a|z, \theta)}{\partial \theta} da < 0$$

$$\frac{\partial^2 U(s, z; \theta, \cdot)}{\partial s \partial z} \quad (13)$$

$$= u''(c_1(\theta)) - \beta \int_{\underline{a}}^{\bar{a}} \frac{\partial y(s, a)}{\partial s} \left[ \Psi' \left( \frac{y(s, a)}{a} \right) \frac{1}{a^2} + \Psi'' \left( \frac{y(s, a)}{a} \right) \frac{y(s, a)}{a^2} \right] \frac{\partial G(a|z, \theta)}{\partial z} da < 0$$

$$\frac{\partial^2 U(s, z; \theta, \cdot)}{\partial z \partial \theta} = -\beta \int_{\underline{a}}^{\bar{a}} \frac{\partial v_2(a, s)}{\partial a} \frac{\partial^2 G(a|z, \theta)}{\partial z \partial \theta} da > 0, \quad (14)$$

applying the envelope theorem and integrating by parts several times, as well as all three assumptions on the conditional distribution function. Also note that  $\frac{\partial y(s, a)}{\partial s} < 0$  because of the income/wealth effect on labor supply.

## A.2 Proof of Lemma 3.1

Consider some admissible reporting strategy  $r(\theta) = \theta'$ .

$$\begin{aligned} \frac{\partial U(\theta, \theta')}{\partial r(\theta)} &= u_c(c_1(\theta')) \frac{\partial c_1(\theta')}{\partial r(\theta)} + \beta \int_{\underline{a}}^{\bar{a}} \frac{\partial v_2(\theta', a)}{\partial r(\theta)} g(a|z(\theta'), \theta) da \\ &\quad + \frac{\partial z(\theta')}{\partial r(\theta)} \int_{\underline{a}}^{\bar{a}} v_2(\theta', a) \frac{\partial g(a|z(\theta'), \theta)}{\partial z(\theta')} da \end{aligned}$$

and

$$\begin{aligned} 0 = \frac{\partial U(\theta', \theta')}{\partial r(\theta)} &= u_c(c_1(\theta')) \frac{\partial c_1(\theta')}{\partial r(\theta)} + \beta \int_{\underline{a}}^{\bar{a}} \frac{\partial v_2(\theta', a)}{\partial r(\theta)} g(a|z(\theta'), \theta') da \\ &\quad + \frac{\partial z(\theta')}{\partial r(\theta)} \int_{\underline{a}}^{\bar{a}} v_2(\theta', a) \frac{\partial g(a|z(\theta'), \theta')}{\partial z(\theta')} da \end{aligned}$$

Subtracting from one another gives:

$$\begin{aligned} \frac{\partial U(\theta, \theta')}{\partial r(\theta)} &= \beta \int_{\underline{a}}^{\bar{a}} \left[ \frac{\partial v_2(\theta', a)}{\partial r(\theta)} (g(a|z(\theta'), \theta) - g(a|z(\theta'), \theta')) \right. \\ &\quad \left. + \frac{\partial z(\theta')}{\partial r(\theta)} v_2(\theta', a) \left( \frac{\partial g(a|z(\theta'), \theta)}{\partial z(\theta')} - \frac{\partial g(a|z(\theta'), \theta')}{\partial z(\theta')} \right) \right] da. \end{aligned}$$

We are now looking when this last expression always has the same sign as the difference  $(\theta - \theta')$ , which is clearly sufficient for global incentive compatibility. For  $(\theta - \theta') > 0$ , by Assumption 2, the first line is positive if  $\frac{\partial v_2(\theta', a)}{\partial r(\theta)}$  or equivalently  $\frac{\partial v_2(\theta, a)}{\partial \theta}$  in a truthful mechanism is increasing in  $a$ . This can be shown to be equivalent to  $\frac{\partial y(\theta, a)}{\partial \theta} > 0$  using the envelope theorem. The second line is positive if  $\frac{\partial z(\theta')}{\partial r(\theta)} > 0$  or equivalently  $\frac{\partial z(\theta)}{\partial \theta} > 0$  in a truthful mechanism. The last condition (iii) is a routine exercise and a proof can be found, for example in Salanié (2003).

### A.3 Second Best Policies: Pareto Problem and Optimality Conditions

After integrating by parts and using the transversality conditions  $\eta(\underline{\theta}) = \eta(\bar{\theta}) = 0$  as well as  $\mu(\theta, \underline{a}) = \mu(\theta, \bar{a}) = 0 \quad \forall \theta$ , the Lagrangian for the social planner's problem reads as

$$\begin{aligned}
& \max_{c_1(\theta), v_2(\theta, a), z(\theta), y(\theta, a)} \int_{\underline{\theta}}^{\bar{\theta}} u(c_1(\theta)) d\tilde{F}(\theta) \\
& + \beta \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) dG(a|z(\theta), \theta) d\tilde{F}(\theta) \\
& + \frac{1}{R} \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{a}}^{\bar{a}} y(\theta, a) dG(a|z(\theta), \theta) dF(\theta) \\
& - \frac{1}{R} \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{a}}^{\bar{a}} u^{-1} [v_2(\theta, a) + \Psi(y(\theta, a)/a)] dG(a|z(\theta), \theta) dF(\theta) \\
& - \lambda_R \int_{\underline{\theta}}^{\bar{\theta}} (c_1(\theta) + z(\theta)) dF(\theta) \\
& - \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{a}}^{\bar{a}} \left( \mu'(\theta, a) v_2(\theta, a) + \mu(\theta, a) \Psi' \left( \frac{y(\theta, a)}{a} \right) \frac{y(\theta, a)}{a^2} \right) da d\theta \\
& - \int_{\underline{\theta}}^{\bar{\theta}} \eta'(\theta) \left[ u(c_1(\theta)) + \beta \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) dG(a|z(\theta)) da \right] d\theta \\
& - \beta \int_{\underline{\theta}}^{\bar{\theta}} \eta(\theta) \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) \frac{\partial g(a|z(\theta), \theta)}{\partial \theta} da d\theta
\end{aligned}$$

With first-order conditions:

$$u'(c_1(\theta))(\tilde{f}(\theta) - \eta'(\theta)) - \lambda_R f(\theta) = 0 \quad (c_{1SB})$$

$$\begin{aligned}
& (\tilde{f}(\theta) - \eta'(\theta)) \beta g(a|z(\theta), \theta) - \lambda_R \frac{1}{R} \frac{1}{u'(c_2(\theta, a))} g(a|z(\theta), \theta) f(\theta) - \mu'(\theta, a) \\
& - \beta \frac{\partial g(a|z(\theta), \theta)}{\partial \theta} \eta(\theta) = 0 \quad (v_{2SB})
\end{aligned}$$

$$\begin{aligned} & \frac{1}{R} \lambda_{Rg}(a|z(\theta), \theta) f(\theta) - \mu(\theta, a) \left[ \Psi'' \left( \frac{y(\theta, a)}{a} \right) \frac{y(\theta, a)}{a^3} + \frac{1}{a^2} \Psi' \left( \frac{y(\theta, a)}{a} \right) \right] \\ & - \frac{1}{R} \lambda_{Rg}(a|z(\theta), \theta) f(\theta) \frac{\Psi' \left( \frac{y(\theta, a)}{a} \right)}{au'(c_2(\theta, a))} = 0, \end{aligned} \quad (y_{SB})$$

$$\begin{aligned} & \tilde{f}(\theta) \beta \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) \frac{\partial g(a|z(\theta), \theta)}{\partial z(\theta)} da + \frac{1}{R} \lambda_{Rf}(\theta) \int_{\underline{a}}^{\bar{a}} \frac{\partial g(a|z(\theta), \theta)}{\partial z(\theta)} (y(\theta, a) - c_2(\theta, a)) da \\ & - \eta'(\theta) \beta \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) \frac{\partial g(a|z(\theta), \theta)}{\partial z(\theta)} da - \beta \eta(\theta) \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) \frac{\partial^2 g(a|z(\theta), \theta)}{\partial z(\theta) \partial \theta} da - \lambda_{Rf}(\theta) = 0 \quad (z_{SB}) \end{aligned}$$

## A.4 Proofs of Propositions 3.4-3.3

### A.4.1 Proposition 3.4

Combining equations  $(c_{1SB})$  and  $(v_{2SB})$  and integrating directly gives the result.

### A.4.2 Proposition 3.2

Rewriting  $(y_{SB})$ :

$$\begin{aligned} & \lambda_{Rg}(a|z(\theta), \theta) f(\theta) \left[ 1 - \frac{\Psi' \left( \frac{y(\theta, a)}{a} \right)}{au'(c_1(\theta, a))} \right] \\ & - R\mu(\theta, a) \left[ \Psi'' \left( \frac{y(\theta, a)}{a} \right) \frac{y(\theta, a)}{a^3} + \frac{1}{a^2} \Psi' \left( \frac{y(\theta, a)}{a} \right) \right] = 0. \end{aligned}$$

Dividing by  $\frac{\Psi'}{au'}$  and  $\lambda_{Rg}(a|z, \theta) f(\theta)$  and using the definition of the labor wedge, i.e.  $u'(1 - \tau_y) = \Psi' \frac{1}{a}$  yields

$$\frac{\tau_y(\theta, a)}{1 - \tau(\theta, a)} = R \frac{\mu(\theta, a)}{\lambda_{Rg}(a|z(\theta), \theta) f(\theta) a} \left[ \frac{\Psi'' \frac{y}{a^2} + \Psi' \frac{1}{a}}{\frac{\Psi'}{au'}} \right],$$

which can be written as

$$\frac{\tau_y(\theta, a)}{1 - \tau(\theta, a)} = R \cdot \frac{\mu(\theta, a)}{\lambda_{Rg}(a|z(\theta), \theta) f(\theta) a} \frac{1 + \varepsilon_u(\theta, a)}{\varepsilon_c(\theta, a)},$$



where  $\frac{\Psi' \frac{1}{a}}{\Psi'' \frac{y}{a^2} + \Psi' \frac{1}{a}} = \frac{1 + \varepsilon_u(\theta, a)}{\varepsilon_c(\theta, a)}$  can be shown by simple algebra, see Saez (2001, p.227). In particular, with the isoelastic specification used in the computations  $\frac{(y/a)^\sigma}{\sigma}$  one can verify that this term is equal to  $\frac{1}{\sigma}$ .

The multiplier  $\mu(\theta, a)$  can be obtained using  $(v_{2SB})$  and  $(c_{1SB})$ :

$$\begin{aligned} \mu(\theta, a) = & \frac{\lambda_R f(\theta)}{u'(c_1(\theta))} \beta G(a|z(\theta), \theta) - \frac{\lambda_R}{R} f(\theta) \int_{\underline{a}}^a \frac{1}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta) \\ & - \beta \frac{\partial G(a|z(\theta), \theta)}{\partial \theta} \eta(\theta^*), \end{aligned}$$

yielding:

$$\frac{\tau_y(\theta, a)}{1 - \tau(\theta, a)} = \frac{1 + \varepsilon_u(\theta, a)}{\varepsilon_c(\theta, a)} \frac{u'(c_2(\theta, a))}{ag(a|z(\theta), \theta)} [\mathcal{A}(\theta, a) + \mathcal{B}(\theta, a)]$$

where

$$\mathcal{A}(\theta, a) = \frac{\beta R G(a|z(\theta), \theta)}{u'(c_1(\theta))} - \int_{\underline{a}}^a \frac{1}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta)$$

$$\mathcal{B}(\theta, a) = -\frac{1}{f(\theta)\lambda_R} \beta R \frac{\partial G(a|z(\theta), \theta)}{\partial \theta} \eta(\theta).$$

Using the inverse Euler equation, the term  $\mathcal{A}(\theta, a)$  can be written as in Proposition 3.4.

From  $(c_{1SB})$ ,  $\eta(\theta)$  is given by:

$$\eta(\theta) = \tilde{F}(\theta) - \lambda_R \int_{\underline{\theta}}^{\theta} \frac{1}{u'(c_1(\theta))} f(\theta) d\theta.$$

The direct benefit of raising utils for agents with skill lower than  $\theta$  is  $\tilde{F}(\theta)$ . The monetary cost is  $\int_{\underline{\theta}}^{\theta} \frac{1}{u'(c_1(\theta))} f(\theta) d\theta$ , transformed into welfare units by  $\lambda_R$ .

### Relation to the formula of Saez (2001)

The insurance part of the labor wedge can be expressed as in Saez (2001), for our case with separable preferences. This relation applies if agents do not differ ex-ante. By some abuse of notation, then  $\mathcal{B}(\theta, a) = 0$  and for  $\mathcal{A}(\theta, a)$ , using the inverse Euler equation, we obtain

$$\begin{aligned}
\mathcal{A}(\theta, a) &= \int_{\underline{a}}^{\bar{a}} \frac{G(a|z(\theta), \theta)}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta) - \int_{\underline{a}}^a \frac{1}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta) \\
&= \int_{\underline{a}}^{\bar{a}} \frac{G(a|z(\theta), \theta)}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta) - \int_{\underline{a}}^{\bar{a}} \frac{1}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta) \\
&\quad + \int_{\underline{a}}^{\bar{a}} \frac{1}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta) \\
&= \int_{\underline{a}}^{\bar{a}} \frac{1}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta) - \int_{\underline{a}}^{\bar{a}} \frac{1 - G(a|z(\theta), \theta)}{u'(c_2(\theta, a^*))} dG(a^*|z(\theta), \theta)
\end{aligned}$$

where the second equality follows from the transversality condition. This term can be expressed as in Saez (2001) as shown by Mankiw, Weinzierl and Yagan (2009) in their online appendix.

### A.4.3 Proposition 3.3

Plugging  $c_{1SB}$  into  $z_{SB}$  gives:

$$\begin{aligned}
&\frac{\lambda_R f(\theta)}{u'(c_1(\theta))} \beta \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) \frac{\partial g(a|z(\theta), \theta)}{\partial z(\theta)} da + \frac{1}{R} \lambda_R f(\theta) \int_{\underline{a}}^{\bar{a}} \frac{\partial g(a|z(\theta), \theta)}{\partial z(\theta)} (y(\theta, a) - c_2(\theta, a)) da \\
&- \beta \eta(\theta) \int_{\underline{a}}^{\bar{a}} v_2(\theta, a) \frac{\partial^2 g(a|z(\theta), \theta)}{\partial z(\theta) \partial \theta} da - \lambda_R f(\theta) = 0
\end{aligned}$$

Proposition 3.3 directly follows.

## A.5 Proof of Proposition 4.1

Starting from a direct mechanism we show in four steps that optimal allocations can indeed be implemented with the policy instruments as defined in Proposition A.5. This idea strongly builds upon the work of Werning (2011).

### Step 1: Introduce savings

Imagine the constrained efficient allocation is implemented by a direct mechanism. From that point on, assume that individuals could freely save  $s$ . Let  $r_1$  denote the report about  $\theta$ . With savings tax  $T^s(s, r_1)$ , the budget constraints read as

$$\tilde{c}_1(r_1) + s = c_1(r_1) \tag{15}$$

$$\tilde{c}_2(r_1, r_2) = c_2(r_1, r_2) + Rs - T^s(s, r_1) \quad (16)$$

Define the optimal report  $r_2$  about  $a$ , for a given report  $r_1$  about  $\theta$ , a given savings tax schedule  $T^s(s, r_1)$  and a given level of savings  $s$ :

$$r_2^*(a, r_1, s, T^s) = \arg \max_{r_2} u [c_2(r_1, r_2) + Rs - T^s(s, r_1)] - \psi \left( \frac{y(r_1, r_2)}{a} \right)$$

Then the optimal report in period one, for a given level of savings and a given savings tax schedule  $T^s(s, r_1)$ , is defined by

$$\begin{aligned} r_1^*(\theta, s, T^s(r_1^1, s)) &= \arg \max_{r_1} u(c_1(r_1) - s) \\ &+ \beta \int_{\underline{a}}^{\bar{a}} u [c_2(r_1, r_2^*) + Rs - T^s(s, r_1)] - \psi \left( \frac{y(r_1, r_2^*)}{a} \right) dG(a|z(r_1^1), \theta) \end{aligned} \quad (17)$$

Then define a hypothetical tax schedule  $T^*(r_1, s, \theta)$  for each  $\theta$  implicitly by

$$V(\theta) = V(\theta, s, r_1^*, T^*(r_1, s, \theta)) \quad \forall s.$$

This hypothetical tax schedule would make individuals of type  $\theta$  indifferent between truth telling and the optimal joint deviation for any  $s$ . It is hypothetical since it does not only depend on the report  $r_1$ , which is observable but also on the unobservable type  $\theta$ . However, we know that for each  $\theta$  such a tax schedule exists. Therefore taking the upper envelope over these functions yields a savings tax function  $\hat{T}(s, r_1)$  that also implements zero savings:

$$\hat{T}(s, r_1) = \sup_{\theta} T^*(r_1, s, \theta) \quad (18)$$

**Lemma A.1.** *A constrained efficient allocation can be implemented by a direct mechanism extended by a savings choice and history-dependent savings tax schedules  $\hat{T}^s(s, r_1)$ .*

## Step 2: Make the savings tax history-independent

A simple way to make the savings tax history-independent is to take the upper envelope of all functions  $T^s(s, r_1)$ , i.e.

$$T^s(s) = \sup_{r_1} \hat{T}^s(s, r_1). \quad (19)$$

**Lemma A.2.** *A constrained efficient allocation can be implemented by a direct mechanism extended by a savings choice and a history-independent savings tax schedule  $T^s(s)$ .*

Note that this savings tax function  $T^s(s)$  is not differentiable and implies zero savings. As Werning (2011) shows one can, using Ricardian equivalence arguments, also construct a history-independent savings tax function that is differentiable and yields non-zero savings choices.

### Step 3: Allow for labor-leisure decisions

To get closer to a decentralized implementation now assume the following extended direct mechanism.

1. Individuals report  $r(\theta)$
2. They get assigned 'income to consume'  $c_1(\theta)$
3. They face the savings tax schedule  $T^s(s)$  and save  $s(\theta) = 0$
4. In period two, instead of directly revealing their type, individuals of type  $\theta$  face an income tax schedule that is defined by

$$T(y(\theta, a), \theta) = y(\theta, a) - c_2(\theta, a) \quad \forall a$$

By the same arguments as in the standard Mirrlees model it follows that this extended direct mechanism can also implement the constrained efficient allocations.

We can summarize this in the following lemma.

**Lemma A.3.** *A constrained efficient allocation can be implemented by a direct mechanism in the first period extended by a savings choice and a history-independent differentiable savings tax schedule  $\tilde{T}^s(s)$  and a history-dependent labor income tax schedule  $T(Y, z)$  in period two.*

### Step 4: Complete Decentralization - allow for educational investment

1. Individuals buy (or tell the government that they want to buy)  $z(\theta)$  units of education

2. They get assigned a student loan  $\mathcal{G}(z(\theta)) = c_1(\theta) + z(\theta)$  (and are obliged to actually buy  $z(\theta)$  units of education)
3. They face the savings tax schedule  $T^s(s)$  and save  $s(\theta) = 0$
4. in period two, instead of directly revealing their type, individuals of type  $\theta$  face an income tax schedule that is defined by

$$T(y(\theta, a), z(\theta)) = y(\theta, a) - c_2(\theta, a) \forall a$$

Since the mechanism in step 4 is just a reformulation of the mechanism in step 3 this directly leads us to Proposition 4.1.

## A.6 Proof of Corollary 4.2

Once  $\theta$  is revealed, it is a direct consequence of Werning (2007, Appendix IV) that there exists a tax system  $T_t(y^t)$  for every  $\theta$  that leads to truth telling of  $a$  conditional on  $\theta$  being revealed. It then directly follows that a capital tax as defined in steps 2 of Appendix A.5 combined with a grant schedule  $\mathcal{G}(z)$  with  $\mathcal{G}(z(\theta)) = z(\theta) + c_e(\theta) \sum_{t=0}^{T_e} \beta^t$  implements the optimum.

## A.7 Graphs from Section 5

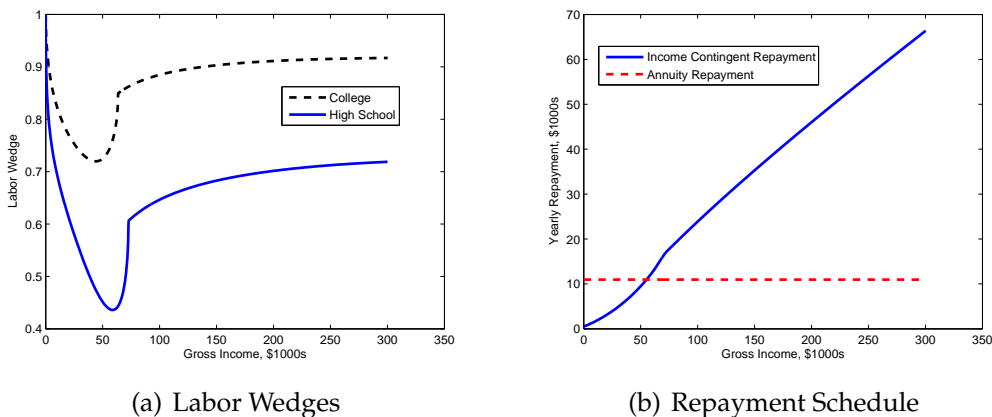
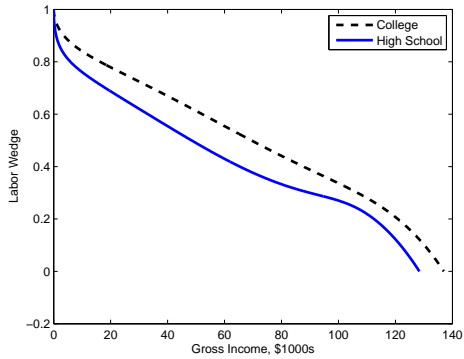
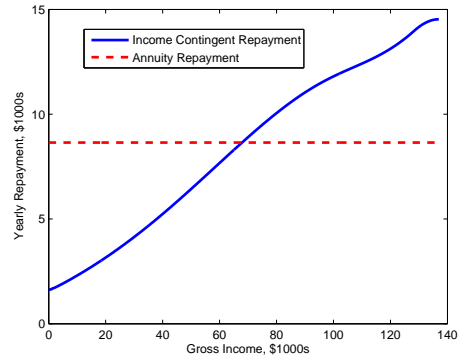


Figure 7: Utilitarian Optimum With Type Dependent Tails

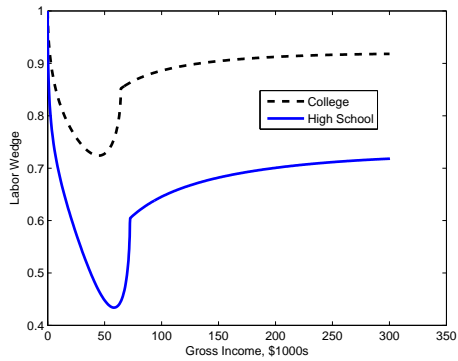


(a) Labor Wedges

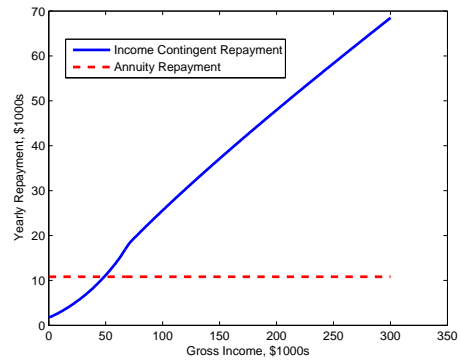


(b) Repayment Schedule

Figure 8: Utilitarian Optimum With Thin Tails

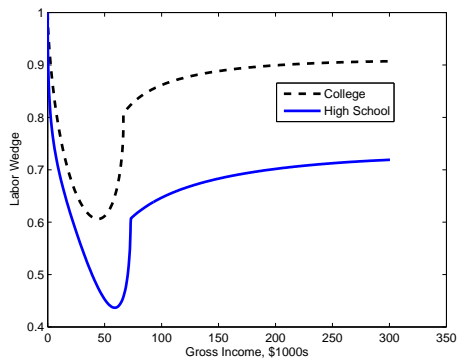


(a) Labor Wedges

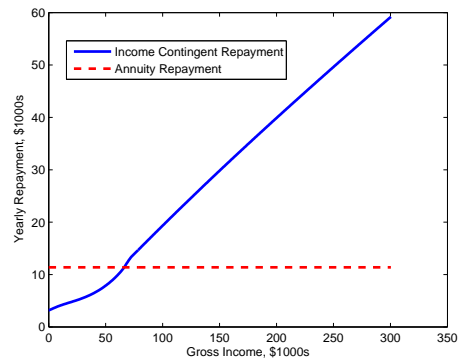


(b) Repayment Schedule

Figure 9: Rawlsian Optimum



(a) Labor Wedges



(b) Repayment Schedule

Figure 10: All Weight on College-Type