Abstract

We study the optimal design of integrated education finance and tax systems. The distribution of wages is endogenously determined by the costly education decisions of heterogeneous individuals before labor market entry. Consistent with empirical evidence, this human capital investment decision is risky. We find that an integrated education and tax system in which the government provides education loans to young individuals coupled with income-contingent repayment can always be designed in a Pareto optimal way. We present a simple empirically driven application of the framework to US data in which individuals make a college entry decision. We find the optimal repayment schemes for college loans can be well approximated by a schedule that is linearly increasing in income up to a threshold and constant afterwards. So although the full optimum could lead to complicated non-linear schedules in theory, very simple instruments can replicate it fairly well. The welfare gains from income-contingent repayment are significant.

JEL-classification: H21, H23, I21

Keywords: Optimal taxation, Education, Education Finance, Student Loans
1 Introduction

How should governments design their higher education finance systems? There exist large differences across countries in the structure of higher education finance. In some countries, such as Denmark, Finland and Sweden, university and college students pay low or no tuition fees and in addition receive grants because of generous public subsidies for higher education. These countries have highly progressive tax systems, which allow to finance these education subsidies. By contrast, in the United Kingdom and the United States, e.g., the burden of educational costs mainly lies on the student and higher education is much less heavily subsidized by public finances. Instead, student loans offered by both the private and the public sector play a big part in financing higher education. From a policy perspective, the choice of an optimal education finance system is intimately linked to the tax system. Both underlie the same basic trade-offs, namely equity concerns in the form of redistribution and insurance against income risk versus efficiency concerns by distorting labor supply and education incentives.

In this paper, we address the optimal design of integrated education finance and tax systems. We build a novel optimal taxation framework in the spirit of Mirrlees (1971) and the vast literature following his footsteps, which allows to study the question from a new angle. In our framework, the distribution of wages is not exogenous but determined by the costly education decisions of individuals before labor market entry. Consistent with what is typically found in empirical studies, this human capital investment decision is risky. To solve the problem, we use an applied mechanism design approach. The benevolent government can observe total income and the education level of individuals, but it has to respect incentive compatibility – first, when individuals decide on education and second, when individuals decide on labor supply. The main novelty of our approach is that in our framework the government is not restricted to the use of predetermined instruments but is free to choose its own instruments, which can condition on education, income and savings. In addition, they are allowed to be fully nonlinear.

We find that an integrated education and tax system in which the government provides education loans to young individuals, coupled with income-contingent repayment rates of these loans after individuals enter the labor market, can effectively deal with all the major trade-offs underlying the education finance and tax problem. In other words, such systems can always be designed such that they are second-best Pareto efficient. This is because income-contingent repayment rates allow the government to effectively differentiate tax distortions across education groups, minimizing the efficiency cost of labor supply distortions. At the same time, it can subsidize education by varying the generosity of the loans.\footnote{We do not model credit market imperfection in the form of borrowing constraints. If these are relevant, as is still a debated question in the literature (Carneiro and Heckman, 2005), wide availability of student loans has the additional benefit of lifting these constraints.} Importantly, the government typically will find it optimal that some individuals partially default
and never pay back the full value of their loans, while for some individuals the amount of repayment might exceed their loan values because this provides insurance.

We present a simple empirically driven application of the framework to US data in which individuals make a college entry decision. We simulate optimal income taxes and college student loans with income-contingent repayment. The optimal policy simulation provides three important insights. First, we find that the optimal repayment scheme for college loans can be well approximated by a schedule that is linearly increasing in income. So although the full optimum could lead to complicated nonlinear schedules in theory, very simple instruments can replicate it fairly well. Second, for our benchmark parameterization college graduates find it optimal to participate voluntarily in the loan schemes as compared to taking a risk-free loan on the private market. Third, we calculate the welfare gains of moving from a third-best scenario where the government optimally sets the income tax and offers a loan system with non-contingent repayment to the system with contingent repayments. We find welfare gains ranging from about 0.2% to 0.6% of lifetime consumption and we show how these gains vary with risk-aversion.

Several countries like the United Kingdom, Australia and New Zealand currently administer income-contingent college student loans, where repayment is proportional to income. Our framework gives these policies a theoretical second-best foundation, based on an applied mechanism design approach to the education finance and taxation problem. Our theoretical considerations suggest that it might be optimal for the government to enforce that very rich individuals pay back more than the capitalized loan value or that repayment might actually be decreasing in income. In the mentioned countries, repayment never exceeds the loan value and repayment schedules are non-decreasing in income. To address these issues, we also consider policy experiments in which we restrict income-contingent repayment not to exceed the actual loan value and to be non-decreasing in income. We find that a large share of the welfare gains from the full optimum can be reaped with these simpler policies.

Relation To Existing Literature. This paper makes a contribution to the literature on optimal income taxation starting from Mirrlees (1971) (see the recent survey of Piketty and Saez (2013)). In Section 3 we discuss how the expression for optimal education-dependent marginal tax rates compares to the seminal optimal tax formulas from Diamond (1998) and Saez (2001) with exogenous human capital.

In two papers, Bovenberg and Jacobs (2005) and Jacobs and Bovenberg (2011) analyze how endogenous education alters the optimal tax problem and discuss to what extent education should be subsidized. Bohacek and Kapicka (2008) study a dynamic model with certainty and obtain equivalent results regarding education subsidies. These articles work under certainty whereas we take idiosyncratic human capital risk into account. Using an...

2Chapman (2006) provides a survey for practices in those and other countries. Barr (2004) discusses the trade-offs involved in designing these programs. To the best of our knowledge, the first economist to endorse the idea was Milton Friedman (1955). He envisioned repayment amounts to be proportional to income, i.e. a linearly increasing repayment schedule. Something we find as an optimal policy in our simulation for the most part of the income distribution.
alytical results and numerical illustrations, we discuss in detail in Section 3 how our findings for optimal education subsidies in a general risky environment relate to their findings. Importantly, with idiosyncratic education risk, the necessity of education dependent labor wedges and income-contingent loans arises, as intuitively they can be understood as providing an additional source of insurance. As we discuss in Section 2, when we review some stylized empirical facts, there is strong evidence that uncertainty about college returns is important and matters for human capital investment decisions.3

Two recent papers, Best and Kleven (2013) and Kapicka and Neira (2013), study how human capital acquisition at the working age influence the optimal taxation problem. We focus on a different part of the human capital accumulation process, namely education before labor market entry. Importantly, both papers reasonably assume that tax policies cannot directly condition on human capital acquired while working. In contrast, we allow the government to use information about education before labor market entry in the tax code, as is done in the real world in some countries in the form of student loans with income-contingent repayment. In addition, our focus is on education finance instead of only tax policies. Working with a two-type model, Gary-Bobo and Trannoy (2013) come to a similar conclusion concerning the income-contingency of loans in a very recent paper. In contrast to their work, we employ a continuous type approach with continuous skill and income distributions in the tradition of the large literature on optimal income taxation as in Mirrlees (1971) and Saez (2001). In particular, we are interested in determining the forces shaping the optimal design of student loan policies both theoretically and numerically, which requires a model with continuous types.

Concerning the implementation of history-dependent allocations, this 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. Also in the context of optimal taxation, Scheuer (2012) considers differential taxation of profits and labor income; in our case a comparable logic applies for an endogenous education instead of an occupational choice.

Finally, taking a quantitative approach and working in the Ramsey tradition with simpler but given policy instruments, Krueger and Ludwig (2013) solve for the optimal income tax and education subsides in a rich macro model.

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 and Section 6 concludes.

2 The Model

2.1 Structure

We consider a stripped-down life-cycle model, in which individuals acquire formal education early in their life cycle and work afterwards. 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 a monetary educational investment $e$. Flow utility during education is denoted by $u_e(c_e)$ with $u_e^c > 0, u_{ee} < 0$. It takes $T_e$ periods (years) until education is finished; the yearly education costs are denoted by $e$. To simplify exposition, we assume that all levels of education take the same amount of time, an assumption we relax later in our optimal policy simulations.

As individuals enter the labor market, they draw their labor market ability $a$ from a continuous conditional cdf $G(a|\theta, e)$, which depends on innate ability $\theta$ and education $e$ and has bounded support $[a, \bar{a}]$, with $a \geq 0$. We assume that preferences over consumption and leisure are given by the utility function $u_w(c_w, l)$, where labor effort $l$ is equal $y/a$, so that gross income is $y = a \times l$. We assume that $u_w(\cdot, \cdot)$ obeys the Spence-Mirrlees condition. The working life lasts for $T_w$ periods.

Expected lifetime utility of an individual of type $\theta$ is hence given by

$$\sum_{t=1}^{T_e} \beta^{t-1} u_e(c_e(\theta)) + \int_0^{T_e+T_w} \sum_{t=T_e+1}^{T_e+t} \beta^{t-1} u_w\left(c_w(\theta, a), \frac{y(\theta, a)}{a}\right) dG(a|\theta, e(\theta)), \quad (1)$$

where we assume the allocation within the education and working period to be constant.$^4$

We write $\beta_e = \sum_{t=1}^{T_e} \beta^{t-1}$ and $\beta_w = \sum_{t=T_e+1}^{T_e+T_w} \beta^{t-1}$. The yearly interest rate in the economy is given by $R = \frac{1}{\beta}$.

As equation (1) reveals, we abstract from further shocks to idiosyncratic labor productivity once individuals have entered the labor market. 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 (versus the

$^4$This is akin to the assumption that the first-order conditions of the second-best problem we solve are also sufficient.
role of shocks over the working life) for lifetime inequality, but different approaches have attributed a major role to it.\(^5\)

Nevertheless, we capture many empirical regularities with this specification of the model. First, assuming \(G(a|e, \theta)\) 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.

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) and Hendricks and Schoellman (2012) suggest 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 \(e\) captures the returns to education. Importantly, for most of our results, we do not impose a certain assumption on the pattern of these returns.

Fourth, as long as \(\frac{\partial^2 G(a|e, \theta)}{\partial \theta \partial e} \neq 0\), 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.\(^6\)

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

**Assumption 1** \(G(a|e', \theta) \succeq_{FOSD} G(a|e, \theta) \iff G(a|e', \theta) \leq G(a|e, \theta), \) for all \(e < e'\) and for all \((\theta, a)\).

**Assumption 2** \(G(a|e, \theta') \succeq_{FOSD} G(a|e, \theta) \iff G(a|e, \theta') \leq G(a|e, \theta), \) for all \(\theta < \theta'\) and for all \((e, a)\).

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

These assumptions 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.

\(^5\)In recent work, Huggett et al. (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 Keane and Wolpin (1997) who attribute 90% to heterogeneity realized before labor market entry, while Storesletten et al. (2004) estimate a number of about 50%.

\(^6\)See also Lemieux (2006) for evidence on heterogeneity in returns.
2.2 Definition of 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 wedges. They are equal to implicit marginal tax rates. We are particularly interested in labor and education wedges. We use subscripts to indicate partial derivatives.

**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:

$$
\tau_y(\theta, a) = 1 - \frac{u_w^w(c_w(\theta, a), \frac{y(\theta,a)}{a})}{u_c^w(c_w(\theta, a), \frac{y(\theta,a)}{a})} \frac{1}{a}.
$$

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

$$
\tau_e(\theta) = 1 - \frac{\beta^w \int_a^\theta u(c_w(\theta, a), \frac{y(\theta,a)}{a}) \frac{\partial g(a|e(\theta), \theta)}{\partial e(\theta)} da}{\beta^e u_e^e(c_e(\theta))}.
$$

Finally, we will also look at optimal distortions of an individual’s Euler equation between the education and the working period.

**Savings Wedge:**

$$
\tau_s(\theta) = 1 - \frac{\beta^w \int_a^\theta u(c_w(\theta, a), \frac{y(\theta,a)}{a}) g(a|e, \theta) da}{\beta^w R \int_a^\theta u(c_w(\theta, a), \frac{y(\theta,a)}{a}) g(a|e, \theta) da},
$$

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

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 $\theta$ at the education stage and $a$ in the working stage. 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 mechanism – agents report an initial type $\theta$ in the education period and, after uncertainty has materialized, report their productivity $a$ in the working period. The planner assigns initial consumption levels $c_e(\theta)$ and education levels $e(\theta)$ to individuals with innate ability $\theta$. Moreover, with each report there comes a sequence of utility promises for the next period $\{v_w(\theta, a)\}_{a \in [a, a]}$. In the second period, the screening takes place over consumption levels $c_w(\theta, a)$ 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 Working 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 $r_\theta(\theta) = \theta$, albeit this is not necessary and just simplifies the exposition.\(^7\) 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 $r_a(a)$:

$$u^w\left(c_w(\theta, a), \frac{y(\theta, a)}{a}\right) \geq u^w\left(c_w(\theta, r_a(a)), \frac{y(\theta, r_a(a))}{a}\right) \quad \forall a, r_a(a), \theta.$$  

Let $v^w(\theta, a)$ be the associated value function. Like in a standard Mirrleesian problem, we assume that 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:

$$\frac{\partial v^w(\theta, a)}{\partial a} = u^w\left(c_w(\theta, a), \frac{y(\theta, a)}{a}\right) \frac{y(\theta, a)}{a^2} \quad (2)$$

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

$$\frac{\partial y(\theta, a)}{\partial a} \geq 0. \quad (3)$$

3.1.2 Education Period Incentive Compatibility

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

$$\beta^u c_e(\theta) + \beta^w \int_a^a v^w(\theta, a)dG(a|e(\theta), \theta) \geq$$

$$\beta^u c_e(\theta) + \beta^w \int_a^\pi v^w(r_\theta(\theta), a)dG(a|e(r_\theta(\theta)), \theta) \quad \forall \theta, r_\theta(\theta).$$

\(^7\)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 $\theta$. 

7
Let $V(\theta)$ be the associated value function. By using the FOC of an agent’s reporting problem, one can easily derive the following envelope condition

$$\frac{dV(\theta)}{d\theta} = \beta^w \int_{a}^{\pi} v^w(\theta, a) \frac{\partial g(a|e(\theta), \theta)}{\partial \theta} da.$$ (4)

As often done in screening problems, our strategy for solving the second-best problem is to work with a relaxed problem with only restrictions (2) and (4) 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.\(^8\)

Next, we present a set of sufficient conditions.

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

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

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

then the considered allocation is incentive compatible.

**Proof.** See Appendix A.1. \(\square\)

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

$$\int_{\theta}^{\pi} \left( \beta^e u^e(c_e(\theta)) + \beta^w \int_{a}^{\pi} v^w(\theta, a) dG(a|e(\theta), \theta) \right) d\tilde{F}(\theta)$$

subject to (2), (4) and the resource constraint:

$$\int_{\theta}^{\pi} \left[ \beta^e (c_e(\theta) + e(\theta)) + \beta^w \int_{a}^{\pi} (c_e(\theta, a) - y(\theta, a)) dG(a|e(\theta), \theta) \right] dF(\theta) = 0.$$
We let the planner assign Pareto weights \( \tilde{f}(\theta) \) to individuals, depending on their initial skill level. Any distribution of these weights, which we normalize to satisfy \( \int_{\theta}^{\tilde{f}(\theta)}d\theta = 1 \), corresponds to one point on the Pareto frontier. \( \tilde{F}(\theta) = \int_{\theta}^{\tilde{f}(\theta)}d\theta \) denotes the cumulative Pareto weight. \( \lambda_R \) denotes the multiplier on the resource constraint and \( \eta(\theta) \) the multiplier function of the first-period envelope conditions. The planner uses the same discount rate as all individuals. We now characterize the wedges of second-best Pareto optimal allocations.

### 3.2.1 Labor Distortions

The following proposition characterizes the optimal labor wedge.\(^9\) For expositional reasons, we focus on the case where utility is separable in consumption and labor and show the formula for the general case in the appendix.

**Proposition 3.2.** Suppose preferences are separable of the form \( u(c) - \Psi(l) \) where \( u'' < 0 \) and \( \Psi'' > 0 \) and further that \( u(\cdot) = u^{e}(\cdot) \). At any constrained Pareto optimum, labor wedges satisfy:

\[
\frac{\tau_y(\theta,a)}{1 - \tau_y(\theta,a)} = \frac{1 + \varepsilon_u(\theta,a)}{\varepsilon^c(\theta,a)} \frac{u'(c_w(\theta,a))}{ag(a|e(\theta),\theta)} \left[ A(\theta,a) + B(\theta,a) \right],
\]

where

\[
A(\theta,a) = G(a|e(\theta),\theta) \left[ \int_{\theta}^{\pi} \frac{1}{u'(c_w(\theta,a^*))dG(a^*|e(\theta),\theta)} \left[ \int_{\theta}^{a} \frac{1}{u'(c_w(\theta,a^*))dG(a^*|e(\theta),\theta)} \right] \right],
\]

\[
B(\theta,a) = \frac{1}{\tilde{f}(\theta)\lambda_R} \frac{\partial [1 - G(a|e(\theta),\theta)]}{\partial \theta} \eta(\theta),
\]

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

\[
\eta(\theta) = \tilde{F}(\theta) - \frac{\int_{\theta}^{\tilde{f}(\theta)}f(\theta)d\theta}{\int_{\theta}^{\tilde{f}(\theta)}\frac{1}{u'(c_w(\theta))}f(\theta)d\theta}.
\]

**Proof.** See Appendix A.2.1.

To understand this analytical result and relate it to the literature, first assume that there would be no incentive problem in period one, i.e. \( \theta \) would be observable. In this case, the term \( B(\theta,a) \) would be zero everywhere because \( \eta(\theta) \) would be zero everywhere. Then, the

\(^9\)In a recent paper Golosov et al. (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.
optimal labor wedge schedules for different values of $\theta$ would be the optimal insurance arrangement for the respective $\theta$-type against income risk. In fact, we show that the formula resembles the standard formula of Saez (2001) for $B(\theta, a) = 0$ in Appendix A.2.1. If $\theta$ were observable, it would be an immutable tag. The planner would want to condition optimal insurance arrangements on $\theta$ in an Akerlof (1978) tagging manner. The interpretation would be very standard that optimal effective marginal tax rates are decreasing in the compensated elasticity, typically larger for higher values of risk aversion and that the nonlinear shape of these effective marginal tax rates is to a large extent determined by the respective distribution function $G(a|\cdot, e(\cdot))$.

With $\theta$ being unobservable, the government has to take incentive compatibility in the first period into account. This is captured by term $B(\theta, a)$. First, note that this term is proportional to the respective value of the Lagrangian-multiplier function $\eta(\theta)$. As long as the planner values the utility of low $\theta$-types sufficiently high (i.e. such that $\tilde{F}(\theta)$ is not too low), $\eta(\theta)$ is positive. This is fulfilled for the Utilitarian case, but also for points on the Pareto frontier more in favor of higher $\theta$-types. If, in addition, the rather natural Assumption 2 applies, term $B(\theta, a)$ is unambiguously positive and thus is a force towards higher effective marginal labor income tax rates.

To get an intuitive understanding for this term, it helps to think about a stylized example. Assume that $G(a|e(\theta^*), \theta^*) = 1$ for all $a > a^*$. Thus, given their choice $e$, individuals of type $\theta^*$ have a zero probability of having a larger labor market skill than $a^*$. In contrast, assume that $G(a|e(\theta^*), \theta^* + \varepsilon) < 1$ for $\varepsilon \to 0$. In that case $\frac{\partial [1-G(a|e(\theta^*), \theta^*)]}{\partial \theta} \to \infty$ for all $a > a^*$ and therefore $\tau_y(\theta^*, a) = 1$ for all $a > a^*$. Intuitively, effective marginal tax rates of 100% for individuals of type $(\theta^*, a)$ with $a > a^*$ have no costs as the mass of individuals whose behavior is distorted is equal to zero. At the same time, these high marginal tax rates make it less attractive for the type with ability $\theta^* + \varepsilon$ to mimic the $\theta^*$-type.

In addition note that the education choice does not play a direct role for the results. If there were no education choices in the first period but individuals would just do nothing, Proposition 3.2.1 would be unchanged. Education only has an indirect effect on the value of optimal effective marginal tax rates through its impact on the distribution of skills. This does not imply that the planner does not take into account the adverse effects of labor supply distortions on the education margin. In fact, the social planner takes this into account by subsidizing the education margin as we discuss in the next subsection.

Finally, a no-distortion at the top and bottom result goes through since $B(\theta, \tilde{a}) = B(\theta, a) = A(\theta, \tilde{a}) = A(\theta, a) = 0$.

3.2.2 Education Distortions

The following proposition characterizes optimal education policies.\footnote{More recently tagging is investigated by Cremer et al. (2010), Mankiw and Weinzierl (2010) as well as Weinzierl (2012).}
Proposition 3.3. At any constrained Pareto optimum, the education wedge is given by:

\[ \tau^e(\theta) = \frac{\beta^w}{\beta^e} \int_0^\pi (y(\theta, a) - c_w(\theta, a)) \frac{\partial g(a|e(\theta), \theta)}{\partial e(\theta)} da 
+ \frac{\beta^w}{\beta^e} \eta(\theta) \int_0^\pi \frac{\partial v^w(\theta, a)}{\partial a} \frac{\partial^2 G(a|e(\theta), \theta)}{\partial e(\theta) \partial \theta} da. \]

Proof. See Appendix A.2.2.

The first term captures the expected marginal fiscal gain of an increase in education. One can show, using integration by parts, 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 \( \frac{\partial^2 G(a|e(\theta), \theta)}{\partial e(\theta) \partial \theta} \) is negative and \( \frac{\partial v^w(\theta, a)}{\partial a} \) is positive everywhere by second-period incentive compatibility. Further, as discussed in the previous subsection, \( \eta(\theta) \) is positive along a large part of the Pareto frontier. 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 \((\theta, \theta)\) result holds.\(^{11}\)

3.2.3 Savings Distortions

The characteristics of savings distortions depend on the properties of the utility function \( u^w(c_w, \frac{y}{a}) \). In the case of separable preferences, the well explored inverse Euler equation holds (Diamond and Mirrlees, 1978; Rogerson, 1985; Golosov et al. 2003), making it optimal to tax savings at every initial skill level \( \theta \), improving the ability of the planner to provide labor supply incentives. In general, the sign of the wedge depends on the exact functional form assumption and especially on the interaction of labor effort and the marginal utility of

\(^{11}\)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 disappears in our environment since the returns to labor supply – once uncertainty has materialized – are independent from the education choice.
consumption – see Golosov et al. (2011) for an elaborate discussion of the underlying forces in a dynamic Mirrleesian model.

3.3 Numerical Illustration

In this section we numerically explore our model in an illustrative manner. We consider two skill distributions as our primitives \( G(a|e, \theta) \) 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 case, 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 compensated elasticity of 0.5 and the CRRA coefficient to \( \rho = 2 \).

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 \( e \). Concerning \( \theta \), we assume a uniform distribution within \([0, 1]\).

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

\[
\mu(\theta, e) = 1.7 + 1.5 \theta^{0.5} e^{0.15}.
\]

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 \( \mu \) is 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,

\[
\mu(\theta, e) = 1.5 + \theta + 0.75 e^{0.25}.
\]

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 a uniform increase in \( e \) across agents. Although \( \frac{\partial^2 \mu(\theta, e)}{\partial \theta \partial e} = 0 \) 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 given 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.\footnote{Following Gallipoli et al. (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 et al., 2011). The realized values of $\mu(\theta, e)$ 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 et al. (2009); as them, we set the scale parameter equal to 0.565.}

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.\textsuperscript{13} Darker regions refer to 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 $\theta$.\textsuperscript{14} In the next panel (b), we illustrate the decomposition from Proposition 3.2 into the insurance term and the incentive term by plotting $A(\theta, a)$ and $B(\theta, a)$. The set of insurance effects $A(\theta, a)$ lies above the set of incentive effects $B(\theta, a)$. Still, especially at the beginning of the income distribution incentive effects contribute to higher implicit tax rates. The graph also reveals that these incentive effects are of more importance for higher innate types on average.

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 \textit{implicit} price distortions of the allocation. In this section, we explore two decentralized implementations of constrained Pareto optima. We focus on utility functions $u(c_w, \frac{y}{a})$ with no income effects, i.e. $u(c_w - \Psi(\frac{y}{a}))$ as in Diamond (1998) or Greenwood et al. (1988). We do this for expositional purposes – in an earlier working paper version (Findeisen and Sachs, 2012) we discuss the implementation with income effects.\textsuperscript{15} Additionally, in the main body, we focus on implementations where education $e(\theta)$ is monotone in type and discuss the case where $e(\theta)$ may be non-monotone in Appendix A.3.2; in this case policies are very similar.

4.1 Implementation One: Education Dependent Taxes

The benevolent government offers a set of student grants to the agents. These grants $G$ are conditional on education. In the working period, there is a tax function, which does not only condition on earnings but also on educational investment.

\textbf{Proposition 4.1.} Suppose there are no income effects and education $e(\theta)$ is strictly monotone. Any constrained Pareto optimal allocation can be implemented by a grant schedule $G(e)$, an education dependent income tax $T(y, e)$ and a savings tax $T^s(s)$, where

\begin{itemize}
  \item $G(e(\theta)) = e(\theta) + c_e(\theta)$
  \item $T(y(\theta, a), e(\theta)) = y(\theta, a) - c_w(\theta, a)$
\end{itemize}

\textsuperscript{13}To economize on space we only show the figures for Case (b) here. The graphs for Case (a) turn out to look nearly identical.

\textsuperscript{14}Since low incomes the distortions are strongly increasing in $\theta$, condition (i) of Lemma 3.1 is typically not fulfilled for low $a$.

\textsuperscript{15}In general, there is also the issue of needing history dependent taxes even with constant wages over the working life, see Werning (2007). We are grateful to Bas Jacobs for pointing out that this can be overcome by the assumption of no income effects.
• $T^*(s)$ is defined as in Appendix A.3.1.

Proof. See Appendix A.3.1

Implementation of Savings Wedges: The savings function $T^*(s)$ 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 $G(e(\theta))$ and $T(y(\theta, a), e(\theta))$ with lump-sum transfers and deductibles to arrive at a nonlinear 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 can be found in Werning (2011).

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

Implementation of Education Wedges: 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 nonlinear grant schedule $G(e)$, 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:

$$\tau^e(\theta) = G'(e) - \int_a^{\pi} \frac{u'(c_w(\theta, a))}{u'(c_e(\theta))} g(a|e(\theta), \theta)T_e(y(\theta, a), e(\theta))da.$$  

A positive value of $\tau^e(\theta)$ encourages education at level $\theta$. The incentive for agents to increase their educational attainment comes from: i) An increase in their grant measured by $G'(e)$ and ii) an increase or decrease in their labor income tax burden for all states, i.e. $T_e(y(\theta, a), e(\theta))$.

4.2 Implementation Two: Income-Contingent Loans

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.\(^\text{17}\) Technically, this can be seen as a simple reinterpretation of the previous implementation – we take the tax system

\(^{16}\)Theoretically it could be that $G$ is (partly) decreasing in $e$ if $c_e(\theta)$ is sufficiently decreasing. However, this is rather unlikely and in all our numerical examples we have $c'_e(\theta) > 0$.

\(^{17}\)Diamond and Saez (2011) argue 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.
of the \( \theta \)-type as the common labor income tax schedule and introduce an income-contingent repayment schedule, which conditions on the size of the loan. Together both instruments are sufficient to replicate the optimal labor wedges. Formally we summarize this in the following proposition:

**Proposition 4.2.** Suppose there are no income effects and education \( e(\theta) \) is strictly monotone. Any constrained Pareto optimal allocation can be implemented by a (compulsory) loan schedule \( L(e) \), a loan repayment schedule \( \Gamma(y,L) \), an income tax \( T(y) \) and a savings tax \( T^s(s) \) where

- \( L(e(\theta)) = e(\theta) + c_e(\theta) \)
- \( \Gamma(y(\theta,a),L(e(\theta))) = c_w(\theta,\tilde{a}(\theta,y(\theta,a))) - c_w(\theta,a) \) if \( y \in [y(\theta,a),y(\theta,\omega)] \) and \( \Gamma(y(\theta,a),L(e(\theta))) = y(\theta,a) - c_w(\theta,a) \) otherwise.
- \( T(y) = y - c_w(\theta,\tilde{a}(\theta,y)) \) \( \forall \ y \in [y(\theta,a),y(\theta,\omega)] \) and \( T = 0 \) otherwise
- \( T^s(s) \) is defined as in Appendix A.3.1,

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

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

\[
c_e(\theta) + e(\theta) \leq L(e(\theta))
\]

\[
c_w(\theta,a) \leq y(\theta,a) - T(y(\theta,a)) - \Gamma(y(\theta,a),L(e(\theta))),
\]

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

The similarity to the other implementation is apparent. Using the agents’ optimality conditions, one can show that the education wedge equals

\[
\tau^e(\theta) = L'(e) - \int_a^\pi \frac{u'(c_w(\theta,a))}{u''(c_e(\theta))} g(a|e(\theta),\theta) \Gamma_L(y(\theta,a),L(e(\theta))) \frac{dL(e(\theta))}{de} da,
\]

and the labor wedge equals

\[
\tau^y(\theta,a) = T'(y(\theta,a)) + \Gamma_y(y(\theta,a),L(e(\theta))).
\]

Education wedges are implemented by the nonlinear loans schedule and how repayment varies with education level. The labor wedge is equal to the marginal tax rate and how loan repayment varies with income.

---

18 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 that income taxes do not exist and all 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 \( \theta \)-type, however, seems to be more natural in our view. Especially in our application of the theory in Section 5.
Note that in Proposition 4.1, we assume the loans to be mandatory. In the numerical simulation we check whether this is a restrictive assumption by allowing college graduates to opt out and instead take a loan with fixed repayment, i.e. with a fixed interest rate. For our baseline parameterization, we find that college students participate voluntarily in the government loan system. Finally, notice that we did not impose a cap on repayments so that in theory for some income and education levels, they might exceed the capitalized loan values. In our numerical simulations, we also consider income-contingent repayment policies, which are not allowed to exceed the loans value.

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 levels of innate ability levels, one that refers to high school and one that refers to college. These simplifications enable us to parameterize the model using factual and, importantly, estimated counterfactual income distributions from the empirical labor literature (Cunha and Heckman, 2007, 2008). Further, the simplification has the advantage that it is easy to incorporate foregone earnings as an implicit cost of education.

5.1 Parametrization

Individuals live for 47 years after they graduate from high-school (age 18-65). Afterwards they enter the labor market directly, or decide to go to college and graduate after four years. We label the two innate types $\theta_{HS}$ and $\theta_{CO}$.\(^{19}\) The incentive constraints read as:

\[
\beta^e u(c_e) + \beta^wCO \int_a^\pi v_{CO}(a, \theta_{CO})g(a|CO, \theta_{CO})da \\
\geq \beta^wHS \int_a^\pi v_{HS}(a, \theta_{HS})g(a|HS, \theta_{CO}), \tag{5}
\]

and

\[
\beta^wHS \int_a^\pi v_{HS}(a, \theta_{HS})g(a|HS, \theta_{HS})da \\
\geq \beta^e u(c_e) + \beta^wCO \int_a^\pi v_{CO}(a, \theta_{CO})g(a|CO, \theta_{HS})da, \tag{6}
\]

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 counterfac-

\(^{19}\)We assume that it is a priori optimal that the low type $\theta_{HS}$ chooses the lower educational attainment (high school) and that $\theta_{CO}$ chooses the higher educational attainment (college).
tual ability distributions. The discount factors take into account the different lengths of the periods, i.e. $\beta^c = \sum_{t=1}^{4} \beta^{t-1}$, $\beta^{wCO} = \sum_{t=5}^{47} \beta^{t-1}$ and $\beta^{wHS} = \sum_{t=1}^{47} \beta^{t-1}$. 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. 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. We assume there is an atom of workers equal to five percent for each distribution reflecting unemployment or disability as in Mankiw et al. (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 et al. (2011), we set the annual monetary cost of college education to $11,100, roughly a third of median income in our data. The yearly interest rate is set to 4% and the yearly discount factor $\beta$ to 1/1.04. We work with a CRRA specification and focus on the case with no income effects so that:

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

---

20 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.

21 Saez (2001) has pioneered the approach to calibrate skill distributions from actual income distributions. We employ the same approximation as in the calibration of the full model in Section 3.3.
with $\sigma = 3$, implying a constant labor supply elasticity of 0.5 and set $\rho = 2$. In unreported simulations, we also varied the values for $\sigma$ and $\rho$; the main results do not change.

5.2 Policies in Baseline Case

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 conservatively assume identical tails for both groups, working with a Pareto parameter of 1.5 (Atkinson et al. 2011; Diamond and Saez, 2011).

5.2.1 Second-Best Optimal Policies

Optimal Labor Wedges: Figure 4(a) displays the optimal labor wedges as a function of yearly income up to $160,000. Both schedules follow a U-shaped pattern, reflecting a 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 they are rather unproductive. Second, the inverse hazard rate $1 - G(a|\cdot, \cdot) / g(a|\cdot, \cdot)$ 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 $1 - G(a|\cdot, \cdot) / ag(a|\cdot, \cdot)$ converges to a constant and as a consequence the labor wedges start to converge.

Note that savings are not distorted in our application. As we assume an education period of length zero for the high school type, there is no transition from an education to a working period, where the planner would find it optimal to distort savings for the high school type. For the college type, we get a no distortion at the top result for savings and therefore, for him the Euler equation holds between the education and the working period.
Looking at Figure 3, one can see in which way tax distortions are tailored to the different income 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. In the top income tails, the wedges converge to almost the same top tax rate (Saez, 2001), with a very small difference caused by the education incentive force $B(\theta, a)$, which we discussed in the theoretical section of this paper, that leads to slightly higher top tax rates for high school types to increase the attractiveness of going to college.23

**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 slope of the repayment schedule is given by the difference in the labor wedges as we outlined in the previous section. As the college wedge lies above the high-school wedge, repayment is increasing in income up to incomes of US-$80,000. Repayments for college graduates start at about US-$1,000. Remarkably, in this income region, the repayment schedule of loans is almost linear with a slope of roughly 0.1, because the difference in the labor wedges is almost constant apart from very low incomes. Afterwards, there is a very small decreasing range and the repayment schedule flattens out as the top labor wedges converge. In sum, optimal repayments can be very well approximated by an intercept of US-$1,000, a US-$1,000 increase in repayment for every US-$10,000 earned up to earnings of US-$70,000 and no additional repayments for incomes above that threshold. So although we did not place any restrictions on the shape of the repayment schedule, linearity comes very close to the second-best optimum.

The red dotted horizontal line shows the yearly repayment that would occur if individuals chose a standard loan (with a yearly interest rate of 4%) where the repayment is not contingent on income and they repay the same amount every year. As can be seen, only some individuals pay back more than in the income-contingent case. This is sensitive to the interest rate, however. For 3%, e.g., more individuals would pay back more in the income-contingent case. For 5%, nobody would pay back more.

As discussed in the implementation section, we assume the college loan system to be mandatory. We check if this is a restrictive assumption by allowing college graduates to opt out and instead take a loan with a yearly interest rate of 4% to finance tuition and early consumption. We find that given the choice, individuals would opt into the loan system with income contingent repayment rates. This is also true for an interest rate of 3%. However, this is arguably a strict test of the assumption since it is not clear whether individuals would be able to borrow up to their desired amount and might face a substantial risk premium on their interest rate if they borrow in the private market.

---

23Some of 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.
5.2.2 Real World Policies: Cap on Repayment and Non-Decreasing Repayment

There might be two limitations to the full second-best optimum which could reduce its real world appeal. First, for some (small) range of the income distribution, repayment for college graduates actually exceeds the loan value, as becomes obvious from Figure 4(b). Second, for high earners the repayment schedule actually decreases in income. These properties are likely to go against commonly held normative views, when it comes to the actual implementation of an income-contingent loan system. Indeed, actual income-contingent repayment systems in the UK or Australia are never decreasing and cap repayment at the loans values. To deal with these concerns, we calculate an allocation which can be implemented with a repayment schedule respecting these constraints – i.e. it is never decreasing and capped at the loan value. In this scenario, effective marginal tax rates for college graduates are adjusted so that they are equal to the marginal tax rates for high school graduates as soon as repayment reaches the capitalized loan value. These modified polices still respect incentive compatibility and budget feasibility, of course.24 Figures 5(a) and 5(b) show the resulting labor wedges and the repayment schedule.

By construction, this repayment system is, of course, inferior in welfare terms to the optimal repayment schedule. As we show in the next subsection, this welfare loss is small.

5.2.3 The Welfare Gains From Income-Contingent Repayment

We now aim at quantifying what the potential welfare effects of income-contingency might be and how much of these welfare gains can be obtained by the (ad-hoc) adjusted repayment schedules, which respect a “non-decreasing constraint” and put a cap on repayment.

24More technically, we first adjusted the lump sum element of the common labor income tax schedule such that the government budget constraint holds. In case, the resulting allocation is not incentive compatible, we adjusted the lump sum elements of the labor income tax and the repayment schedule such that the government budget constraint holds and the incentive constraint of the college type binds.
The natural policy comparison is the case where repayment is not contingent on income. For this benchmark case, we allow the government to freely choose an income tax schedule and also optimize over education subsidies and savings taxes. Formally, the only additional restriction is that individuals with the same income should face the same labor wedge.

To be able to make such a welfare comparison, the crucial assumption is the absence of income effects. In this case, the restriction that the labor wedge is only a function of current income is simply equivalent to:

\[ y(\theta, a) = y(a). \]  

(7)

The following proposition states how a Pareto optimal allocation, subject to (7), can be implemented in the binary education model.

**Proposition 5.1.** Assume there are no income effects. Then any Pareto optimal allocation subject to private information and (7) 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

- \( T(y(a)) = y(a) - c(\theta_{HS}, a) \)
- \( \Gamma = c(\theta_{HS}, a) - c(\theta_{CO}, a) \)
- \( L = \beta^e (c_e + e) \).

Figure 6 shows the optimal education independent labor income tax in this case; the optimal marginal tax rates lie between their education dependent counterparts from the second best optimum.

We next calculate the welfare gains from income-contingent repayment schemes for both cases: the unrestricted repayment schedule from Section 5.2.1 and the constrained one from Section 5.2.2.

In Figure 7 we present the consumption equivalent welfare gains as the CRRA parameter \( \rho \) varies from 1 to 4. First, one can see that the “real world appeal”-repayment schedule is able to reap almost all the welfare gains from income-contingent loans. Second, the gains are increasing in risk-aversion which underscores the role of the loans as an insurance device.

---

25 In contrast, in the case with income effects, education-independent marginal tax rates do not imply \( y(\theta, a) = y(a) \). Concretely, as individuals with the same \( a \) but different \( \theta \) typically differ concerning their optimal consumption, they choose different labor effort although they face the same labor wedge schedule. Imposing the same consumption and income level on all individuals with the same skill level \( a \) would overcome this problem, however, it would be a much stronger restriction on optimal policies. All these arguments do not necessarily imply that one cannot compute optimal history independent policies for the case with income effects. For this case, however, we would not be able to use a first-order approach but instead it would be necessary to check all possible incentive constraints. This would require us to significantly reduce the type space, severely limiting our ability to characterize nonlinear schedules and make welfare comparisons across scenarios.

26 Naturally, another implementation of this optimum would involve a single labor tax schedule with education-dependent lump-sums and education grants offered by the government.
For a CRRA coefficient of two, the gains are about 0.32% in the unrestricted and about 0.25% in the restricted case. Thus, roughly 78% of the welfare gain from the second-best can be reaped with the restricted repayment. In case of an interest rate of 3%, 68% of the welfare gain can be reaped with simpler policies. For an interest rate of 5%, second-best optimal income-contingent repayment would actually never exceed the loan value.

Finally, the welfare gains are evenly distributed in the benchmark case ($\rho = 2$), implying that both the college and the high school type achieve a utility gain of 0.32% of consumptions equivalents. For lower values of $\rho$, a larger share of the gain is reaped by the high-school graduates, for higher values of $\rho$ the result is reversed.

### 5.3 Policies in Case of Differing Top Income Tails

We now test if and how a different assumption on top incomes across income distributions changes the results. We focus on the case, where the college income distribution has a thicker tail than the high school income distribution. For college graduates, we choose a Pareto pa-
rameter of 1.28. For high-school graduates we choose a Pareto parameter of 3.27 These values lie within the range of what has been typically found in empirical studies covering many countries and time periods (Atkinson et al, 2011). If we aggregate the two distributions to the aggregate income distribution, we find that the resulting tail for top incomes resembles a Pareto tail with a parameter not far away from 1.5.28

5.3.1 Second-Best Optimal Policies

Figures 8(a) and 8(b) display the corresponding schedules for labor wedges and the repayment schedule. The college labor wedge now lies above the high school labor wedge everywhere, leading to a strictly increasing repayment schedule. The implicit top tax rate for college graduates is higher than for high-school graduates, driven by the differences in the Pareto parameter. Interestingly, again a simple linear approximation of the repayment schedule with a linear slope of about 11% could almost perfectly implement the second-best optimum. Repayment of college graduates now exceeds the annuity loan value by a much more significant amount and for much bigger fraction of the population. We check again if a college type would prefer not to choose the income-contingent loan in this case and find that the loans indeed have to be compulsory. However, as we show next, one can again construct slightly different policies which respect a cap on repayment. These yield a large share of the welfare gain and do not require the loans to be compulsory.

27The top tails are not dependent on innate type $\theta$ but are just determined by the education level. In an earlier working paper version (Findeisen and Sachs, 2012), we also explore the case in which the tails are determined by innate type $\theta$ instead. The results are very similar.

28The sum of two Pareto distributions tends to behave like a Pareto distribution, where the heavier tail distribution seems to dominate (Ramsay, 2006). This implies that, in the tails, the resulting aggregate distribution is very close to the college distribution.
5.3.2 Real World Polices: Cap on Repayment

As in Section 5.2.2, we now adjust the second-best optimum towards policies that satisfy the same two mentioned real-world restrictions. The adjustment we make is slightly different this time. In Section 5.2.2, we lowered the labor wedges of the college types such that they equal the optimal ones for the high school types above all income levels, where the second-best repayment starts to exceed the loan value. Here, we do the opposite and increase the labor wedges of the high school types such that they are equal to the college labor wedges. The reason for this is that optimal history independent wedges (see Figure 10) are closer to the college wedges for high incomes, which is driven by the fatter college top income tail “dominating” the top income tail for the high school types, see footnote 28. The new adjusted policies respect incentive compatibility and budget feasibility. In order to avoid bunching because of a discrete upward jump in marginal tax rates, we smooth out the increase over an interval of roughly US-$5,000. The resulting labor wedges and repayment are illustrated in Figures 9(a) and 9(b).

5.3.3 The Welfare Gains From Income-Contingent Repayment

As in Section 5.2.3, we now calculate the welfare gains over student loans without income-contingent repayment. Due to the differing top income tails, the college and high school wedges are more distinct from each other (see Figure 10) than in the benchmark case. This yields to welfare gains (see Figure 11) that are slightly higher. They are 0.36% of lifetime consumption for a CRRA coefficient of 2. Again, the adjusted system respecting a cap can yield a large part of those gains: in fact, they lead to a gain of 0.33%, which is almost 92% of the welfare gain. For an interest rate of 3% (5%) the latter value is 75% (95%).
Figure 10: Optimal Education Independent Taxes

Figure 11: Welfare Gains and Risk-Aversion

6 Conclusion

This paper has 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. We therefore make a second-best argument in favor of student loans with income-contingent repayment rates and, in addition, provide guidance for the optimal design of such repayment schedules.

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. Closely 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 abil-
ities 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 no consider human capital accumulation after labor market entry like on-the-job training.\textsuperscript{29} Third, we assumed full commitment to all policies from the government side. Relaxing these assumptions might be a fruitful area for future research.

\textsuperscript{29}In ongoing research, Stantcheva (2012) considers optimal taxation and human capital taxation in a life cycle economy, which encompasses on-the-job training.
References


A Appendix

A.1 Proof of Lemma 3.1

Consider some admissible reporting strategy \( r(\theta) = \theta' \). Denote by \( U(\theta, \theta') \) the utility obtained by a type \( \theta \) reporting \( \theta' \). Then consider the following derivatives

\[
\frac{\partial U(\theta, \theta')}{\partial r(\theta)} = u_c(c_e(\theta')) \frac{\partial c_e(\theta')}{\partial r(\theta)} + \beta \int_a^\pi \frac{\partial v^w(\theta', a)}{\partial r(\theta)} g(a|e(\theta'), \theta) da
\]

\[
+ \frac{\partial e(\theta')}{\partial r(\theta)} \int_a^\pi v^w(\theta', a) \frac{\partial g(a|e(\theta'), \theta)}{\partial e(\theta')} da
\]

and

\[
0 = \frac{\partial U(\theta', \theta')}{\partial r(\theta)} = u_c(c_e(\theta')) \frac{\partial c_e(\theta')}{\partial r(\theta)} + \beta \int_a^\pi \frac{\partial v^w(\theta', a)}{\partial r(\theta)} g(a|e(\theta'), \theta') da
\]

\[
+ \frac{\partial e(\theta')}{\partial r(\theta)} \int_a^\pi v^w(\theta', a) \frac{\partial g(a|e(\theta'), \theta)}{\partial e(\theta')} da
\]

which is equal to zero by first-order incentive compatibility. Subtracting from one another gives:

\[
\frac{\partial U(\theta, \theta')}{\partial r(\theta)} = \beta \int_a^\pi \left[ \frac{\partial v^w(\theta', a)}{\partial r(\theta)} \left( g(a|e(\theta'), \theta) - g(a|e(\theta'), \theta') \right) \right. \\
+ \left. \frac{\partial e(\theta')}{\partial r(\theta)} v^w(\theta', a) \left( \frac{\partial g(a|e(\theta'), \theta)}{\partial e(\theta')} - \frac{\partial g(a|e(\theta'), \theta')}{\partial e(\theta')} \right) \right] da.
\]

We are now looking when this last expression always has the same sign as the difference \((\theta - \theta')\), which is sufficient for global incentive compatibility. For \((\theta - \theta') > 0\), using Assumption 2, the first line is positive if \( \frac{\partial v^w(\theta', a)}{\partial r(\theta)} \) (or equivalently \( \frac{\partial v^w(\theta, a)}{\partial \theta} \) in a truthful mechanism) is increasing in \( a \). This can be shown to be equivalent to \( \frac{\partial g(a|e(\theta'), \theta)}{\partial e(\theta')} > 0 \) using the envelope theorem, which is part (i) of the lemma. Using assumption 3, one can show, that the second line is positive if \( \frac{\partial e(\theta')}{\partial r(\theta)} > 0 \) or equivalently \( \frac{\partial e(\theta)}{\partial \theta} > 0 \) in a truthful mechanism, which is part (ii) of the lemma. That (2) and (3) are sufficient for global incentive compatibility in the working period is a routine exercise and a proof can be found, for example in Salanié (2003).

A.2 Optimal Labor and Education wedges

We start by stating the objective for the case of separable preferences of the form \( u(c) - \Psi(l) \), where \( \Psi \) are the convex utility costs of labor. Further we assume that \( u(\cdot) = u(\cdot) \). After
integrating by parts and using the transversality conditions $\eta(\theta) = \eta(\bar{\theta}) = 0$ as well as $\mu(\theta, a) = \mu(\theta, \bar{a}) = 0 \ \forall \theta$, the Lagrangian for the social planner’s problem reads as:

$$\max_{c_e(\theta), v^w(\theta,a), c(\theta), y(\theta,a)} \mathcal{L} = \beta^e \int_\theta^\infty u(c_e(\theta))d\tilde{F}(\theta)$$

$$+ \beta^w \int_\theta^\infty \int_a^\infty v^w(\theta,a)dG(a|c(\theta), \theta)d\tilde{F}(\theta)$$

$$+ \beta^w \lambda_R \int_\theta^\infty \int_a^\infty y(\theta,a)dG(a|c(\theta), \theta)dF(\theta)$$

$$- \beta^w \lambda_R \int_\theta^\infty \int_a^\infty u^{-1}[v^w(\theta,a) + \Psi(y(\theta,a)/a)]dG(a|c(\theta), \theta)dF(\theta)$$

$$- \beta^w \lambda_R \int_\theta^\infty (c_e(\theta) + e(\theta))dF(\theta)$$

$$- \int_\theta^\infty \int_a^\infty (\mu'(\theta,a)v^w(\theta,a) + \mu(\theta,a)\Psi'\left(\frac{y(\theta,a)}{a}\right)\frac{y(\theta,a)}{a^2})d\alpha d\theta$$

$$- \int_\theta^\infty \eta'(\theta)\left[\beta^e u(c_e(\theta)) + \beta^w \int_a^\infty v^w(\theta,a)dG(a|c(\theta))da\right]d\theta$$

$$- \beta^w \int_\theta^\infty \eta(\theta) \int_a^\infty v^w(\theta,a)\frac{\partial g(a|c(\theta), \theta)}{\partial \theta}d\alpha d\theta$$

With first-order conditions:

$$u'(c_e(\theta))(\tilde{f}(\theta) - \eta'(\theta)) - \lambda_R f(\theta) = 0$$

$$\left(\tilde{f}(\theta) - \eta'(\theta)\right)g(a|c(\theta), \theta) - \lambda_R \frac{1}{u'(c_e(\theta), a)}g(a|c(\theta), \theta)f(\theta) - \frac{\mu'(\theta, a)}{\beta^w}$$

$$- \frac{\partial g(a|c(\theta), \theta)}{\partial \theta} \eta(\theta) = 0$$

$$\lambda_R g(a|c(\theta), \theta)f(\theta) - \frac{\mu(\theta,a)}{\beta^w} \left[\Psi'\left(\frac{y(\theta, a)}{a}\right)\frac{y(\theta,a)}{a^3} + \frac{1}{a^2}\Psi'\left(\frac{y(\theta,a)}{a}\right)\right]$$

$$- \lambda_R g(a|c(\theta), \theta)f(\theta) \frac{\Psi'\left(\frac{y(\theta,a)}{a}\right)}{u'(c_e(\theta),a)} = 0$$

\[30\text{With more general preferences, the fourth line would be } \beta^w \lambda_R \int_\theta^\infty \int_a^\infty \gamma(v^w(\theta,a), \frac{y(\theta,a)}{a})dG(a|c(\theta), \theta)dF(\theta) \]

\[\text{with } \gamma(v, l) \text{ being the inverse function of } u \text{ over } c. \text{ The sixth line would be } - \beta^w \lambda_R \int_\theta^\infty (c_e(\theta) + e(\theta))dF(\theta) - \int_\theta^\infty \int_a^\infty \left(\mu'(\theta, a)v^w(\theta,a) + \mu(\theta, a)u_l\left(\gamma\left(v^w(\theta,a), \frac{y(\theta,a)}{a}\right), \frac{y(\theta,a)}{a^2}\right)\right)\frac{u'(\theta,a)}{\beta^w}\right)d\theta \text{ instead.} \]
\[ \dot{f}(\theta) \beta^w \int_0^\pi v^w(\theta, a) \frac{\partial g(a|e(\theta), \theta)}{\partial e(\theta)} da + \beta^w \lambda_R f(\theta) \int_0^\pi \frac{\partial g(a|e(\theta), \theta)}{\partial e(\theta)} (y(\theta, a) - c_w(\theta, a)) da \]

\[ -\eta'(\theta) \beta^w \int_0^\pi v^w(\theta, a) \frac{\partial g(a|e(\theta), \theta)}{\partial e(\theta)} da - \beta^w \eta(\theta) \int_0^\pi v^w(\theta, a) \frac{\partial^2 g(a|e(\theta), \theta)}{\partial e(\theta) \partial \theta} da - \beta^w \lambda_R f(\theta) = 0 \]

Combining equations (8) and (9) and integrating directly gives the inverse euler equation.

A.2.1 Proposition 3.2

Rewriting (10):

\[ \lambda_R g(a|e(\theta), \theta) f(\theta) \left[ 1 - \frac{1}{\beta^w} \lambda_R g(a|e(\theta), \theta) f(\theta) \right] \]

\[ -\frac{1}{\beta^w} \mu(\theta, a) \left[ \Psi'' \left( \frac{y(\theta, a)}{a} \right) \frac{y(\theta, a)}{a^3} + \frac{1}{a^2} \Psi' \left( \frac{y(\theta, a)}{a} \right) \right] = 0. \]

Dividing by \( \psi_{a} \) and \( \lambda_R g(a|e, \theta) f(\theta) \) and using the definition of the labor wedge, i.e. \( u'(1 - \tau_y) = \Psi_{a}^1 \) yields

\[ \frac{\tau_y(\theta, a)}{1 - \tau(\theta, a)} = \frac{1}{\beta^w} \lambda_R g(a|e(\theta), \theta) f(\theta) \left[ \frac{\Psi'' \left( \frac{y(\theta, a)}{a} \right) + \Psi_{a}^1}{\psi_{a}} \right], \]

which can be written as

\[ \frac{\tau_y(\theta, a)}{1 - \tau(\theta, a)} = \frac{1}{\beta^w} \cdot \frac{\mu(\theta, a)}{\lambda_R g(a|e(\theta), \theta) f(\theta) a} \left[ 1 + \epsilon_u(\theta, a) \right] \frac{1}{\epsilon_e(\theta, a)}, \]

where \( \frac{\Psi_{a}^1}{\psi_{a} + \psi_{a}^2} = \frac{1 + \epsilon_u(\theta, a)}{\epsilon_e(\theta, a)} \) can be shown by simple algebra, see Saez (2001, p.227). In particular, with the isoelectric specification used in the computations \( \left( \frac{y}{a} \right)^{\sigma} \) one can verify that this term is equal to \( \frac{1}{\sigma} \).

The multiplier \( \mu(\theta, a) \) can be obtained using (9) and (8):

\[ \frac{\mu(\theta, a)}{\beta^w} = \frac{\lambda_R f(\theta)}{u'(c_e(\theta))} G(a|e(\theta), \theta) - \lambda_R f(\theta) \int_0^a \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta) \]

\[ -\frac{\partial G(a|e(\theta), \theta)}{\partial e(\theta)} \eta(\theta^*), \]

yielding:

\[ \frac{\tau_y(\theta, a)}{1 - \tau(\theta, a)} = \frac{1 + \epsilon_u(\theta, a)}{\epsilon_e(\theta, a)} \frac{u'(c_w(\theta, a))}{u'(c_w(\theta, a))} \left[ A(\theta, a) + B(\theta, a) \right] \]
The direct benefit of raising utils for agents with skill lower than \( \theta \) is given by:

\[
A(\theta, a) = \frac{G(a|e(\theta), \theta)}{u'(c_e(\theta))} - \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta)
\]

\[
B(\theta, a) = -\frac{1}{f(\theta)\lambda_R} \frac{\partial G(a|e(\theta), \theta)}{\partial \theta} \eta(\theta).
\]

Using the inverse Euler equation, the term \( A(\theta, a) \) can be written as in the proposition:

\[
\frac{G(a|e(\theta), \theta)}{u'(c_e(\theta))} - \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta)
= G(a|e(\theta), \theta) \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta) - \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta)
= G(a|e(\theta), \theta) \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta) + G(a|e(\theta), \theta) \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta)
- \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta)
= G(a|e(\theta), \theta) \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta) - (1 - G(a|e(\theta), \theta)) \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta).
\]

From (8), \( \eta(\theta) \) is given by:

\[
\eta(\theta) = \tilde{F}(\theta) - \lambda_R \int_\theta^\pi \frac{1}{u'(c_e(\theta))} f(\theta) d\theta.
\]

The direct benefit of raising utils for agents with skill lower than \( \theta \) is \( \tilde{F}(\theta) \). The monetary cost is \( \int_\theta^\pi \frac{1}{u'(c_e(\theta))} f(\theta) d\theta \), 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 \( B(\theta, a) = 0 \) and for \( A(\theta, a) \), using the inverse Euler equation, we obtain

\[
A(\theta, a) = \int_a^\pi G(a|e(\theta), \theta) \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta) - \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta)
= \int_a^\pi G(a|e(\theta), \theta) \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta) - \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta)
+ \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta)
= \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta) - \int_a^\pi \frac{1}{u'(c_w(\theta, a^*))} dG(a^*|e(\theta), \theta).
\]
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.

**General Preferences:** Carrying out the analogous steps with a general utility function $u(c,l)$ we get:

$$\frac{\tau_y(\theta,a)}{1 - \tau(\theta,a)} = 1 + \epsilon_u(\theta,a) \frac{u_c(\theta,a)}{ag(a|e(\theta),\theta) \int_a^\pi \exp \left( - \int_a^x \frac{u_c(l(\theta,s),l(l(\theta,a))}{u_c(l(\theta,s),\theta)} ds \right) \times [A(\theta,x) + B(\theta,x)] dx,$$

where $A(\theta,x) = g(x|e(\theta),\theta) \left( \frac{1}{w(c_e(\theta,x))} - \frac{1}{w(c_w(\theta,x))} \right)$ and $B(\theta,x) = \eta(\theta) \frac{\partial g(x|e(\theta),\theta)}{\partial \theta}.$

### A.2.2 Proposition 3.3

Plugging (8) into (11) gives:

$$\frac{\lambda_R f(\theta)}{u'(c_e(\theta))} \int_a^\pi v^w(\theta,a) \frac{\partial g(a|e(\theta),\theta)}{\partial e(\theta)} da + \beta^w \lambda_R f(\theta) \int_a^\pi \frac{\partial g(a|e(\theta),\theta)}{\partial e(\theta)} (y(\theta,a) - c_w(\theta,a)) da$$

$$- \beta^w \eta(\theta) \int_a^\pi v^w(\theta,a) \frac{\partial^2 g(a|e(\theta),\theta)}{\partial e(\theta) \partial \theta} da - \beta^e \lambda_R f(\theta) = 0$$

Proposition 3.3 directly follows. Note that the relevant first-order conditions are identical for general utility function, so that the formula for the optimal wedge in the proposition also applies.

### A.3 Implementation Appendix

#### A.3.1 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 4.1. The idea to work with a history independent savings tax 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 $s$ at rate $R$. Let $r_1$ denote the report about $\theta$. With savings tax $T^s(s,r_1)$, the budget constraints read as

$$\tilde{c}_w(r_1) + s = c_w(r_1)$$

(12)

$$\tilde{c}_w(r_1,r_2) = c_w(r_1,r_2) + Rs - T^s(s,r_1)$$

(13)
Define the optimal report $r_2$ about $a$, for a given report $r_1$ about $\theta$, a given savings tax schedule $T^*(s, r_1)$ and a given level of savings $s$:

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

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

$$r_1^*(\theta, s, T^*(r_1, s)) = \arg \max_{r_1} u(c_e(r_1) - s) + \beta \int_a \left[ c_w(r_1, r_2^*) + Rs - T^*(s, r_1) - \psi \left( \frac{y(r_1, r_2^*)}{a} \right) \right] dG(a|e(r_1), \theta) \quad (14)$$

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

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

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 and is feasible since it does not condition on $\theta$:

$$\hat{T}(s, r_1) = \sup_{\theta} T^*(r_1, s, \theta). \quad (15)$$

**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 $\hat{T}(s, r_1)$.

**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 $T^*(s, r_1)$, i.e.

$$T^*(s) = \sup_{r_1} \hat{T}(s, r_1). \quad (16)$$

**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 $T^*(s)$.

Note that this savings tax function $T^*(s)$ 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.

---

31Recall that $V(\theta)$ is the value function of a truth teller of type $\theta$. 

37
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’ \( c_r(\theta) \)
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

\[
T(y(\theta,a), e(\theta)) = y(\theta,a) - c_w(\theta,a) \forall a.
\]

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 \( T(y,e) \) in period two.

Step 4: Complete Decentralization – allow for educational investment

1. Individuals buy (or tell the government that they want to buy) \( e(\theta) \) units of education
2. They get assigned a student loan \( G(e(\theta)) = c_e(\theta) + e(\theta) \) (and are obliged to actually buy \( e(\theta) \) 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

\[
T(y(\theta,a), e(\theta)) = y(\theta,a) - c_w(\theta,a) \forall a
\]

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.3.2 Discussion of Implementations with Non-Monotone Education

If education is not strictly monotone, it is not enough to condition tax and grant schedules on education for education levels which are assigned to more than one type. In this case for those respective education levels, the planner can augment the system of education grants,
such that there is no more a unique grant per education level but a set of grants, which contains the respective correct grant. More formally, if the education level $e^*$ is the optimal education level for all individuals within a certain set $\Theta(e^*)$, then the set of grants assigned to education level $e^*$ must contain $G(e^*, \theta^*) = c_e(\theta^*) + e^*$, for each $\theta^* \in \Theta(e^*)$. Every education dependent tax function associated with a level $e^*$, can then in addition be conditioned on consumption during education $c_e(\theta^*)$. Analogously, the planner can offer multiple loan sizes per education level, and repayment schedules which condition on the loan size, income and early consumption.