KU LEUVEN

CENTER FOR ECONOMIC STUDIES

DISCUSSION PAPER SERIES
DPS15.21

OCTOBER 2015

The economics of financing higher education

Ron DIRIS and Erwin OOGHE

Public Economics

Faculty of Economics
and Business

THE ECONOMICS OF FINANCING HIGHER EDUCATION*

Ron Diris · Erwin Ooghe

October 9, 2015

1 Motivation

Participation in higher education has increased considerably over the last decades. The resulting budgetary pressure, in combination with the recent financial crisis, has led to reforms in the financing schemes for higher education many countries. These reforms often shift a substantial share of the costs towards students and their parents. Not only have tuition fees increased, but new modes of financing—for example, fees and grants that depend on study progress, or loan repayments that are contingent on future income—are increasingly introduced as well.

Participation in higher education is unlikely to decline in the future and the resulting budgetary pressure is therefore likely to stay. In addition, other budgetary needs, in particular the financing of health care and pensions, are higher on the political agenda in many countries. Ad hoc reforms in the financing of higher education can alleviate some of the budgetary needs in the short-run, but there is a great need to design structural policies that are sustainable in the long run as well.

The current review spells out economic arguments to guide our thinking about the financing of higher education. Such a review is not only useful to the current public debate in many countries, it also provides a well-timed update of the current academic literature. First, some of the older arguments, like externalities and signalling, turn out

*This paper has benefited from discussions with André Decoster, Walter Nonneman, Erik Schokkaert, Tom Truyts, Dirk Van Damme, and Frank Vandenbroucke. We also thank audiences at Euroforum, Leuven 2013 and the Canazei 2014 winter school for helpful comments. Ron Diris, Department of Economics, KU Leuven (ron.diris@kuleuven.be); Erwin Ooghe, Department of Economics, KU Leuven and ZEW, Mannheim (erwin.ooghe@kuleuven.be).

to be difficult to validate empirically, while others, e.g., credit constraints, have become more important over time. We review the empirical literature with special attention to the causal evidence that has become increasingly available over the last years. Second, recent insights from behavioural economics provide new arguments for intervention in higher education. Additionally, multiple recent studies in the public finance literature have analyzed the pros and cons of different innovative financing schemes, such as income-contingent loans and graduate taxes. We integrate the theoretical and empirical findings of behavioural economics and public finance in the current review.

The starting point of our analysis in section 2 is the following question: do we need government intervention in higher education? The second welfare theorem provides a clear answer. If there are no market failures, if individuals behave rationally, and if non-distortive lump-sum transfers are feasible, then there is no need for government intervention other than redistribution based on lump-sum transfers.

The obvious next question is: do these assumptions hold? Economic arguments for government intervention may arise if one of the assumptions fail. Insurance, credit, education and labour markets—all highly revelant to higher education—may malfunction in the absence of government intervention. In addition, there is ample evidence that individuals do not always behave rationally, and students turn out to be no exception. Finally, although non-distortive lump-sum transfers are (first) best, they are not feasible in reality.

In sections 3, 4, and 5 we discuss market failures (incomplete insurance, credit constraints, externalities, and signalling), behavioural failures (misprediction, positional goods, and peer interactions), and first-best failures (distortions, perverse redistribution, and fiscal externalities). In each section we identify the theoretical problem, analyze the consequences for participation in higher education, and provide the available empirical evidence.

The final section seeks to answer one last question: how should educational subsidies and redistribution policies look like? In section 6 we first take stock and identify the failures that we consider essential for an optimal policy design. Afterwards, we discuss the pros and cons of the different financing modes (general taxation, classical loans, income-contingent loans, and graduate taxes) in more detail.

We focus on the economics of financing higher education, so other fields (e.g., psychology and educational sciences), other interventions (e.g., regulation), other educational levels (e.g., compulsory education), and other tasks of higher education (e.g., research) only enter the discussion if deemed essential. In addition, the current review is limited to the study of demand side aspects (versus supply and governance). In particular, we mainly focus on the participation decision (versus effort and success) in higher education from a welfare point of view (versus a political economy view).

2 Do we need government intervention in higher education?

We consider the following laissez-faire policy in higher education: students pay the market price of higher education and the government does not intervene except for redistribution based on non-distortive lump-sum transfers. The question is: under what assumptions is the laissez-faire the best possible policy for society? The second welfare theorem provides the answer. In this section we spell out the assumptions underlying the second welfare theorem and show how they justify the laissez-faire. In later sections we will scrutinize the underlying assumptions. To set the stage, it is useful to start out with a review of the expected private costs and benefits of participation in higher education.

2.1 Expected private costs and benefits of participation

Economic theory dictates that rational students participate in higher education if the expected private benefits of participation exceed the private costs. But what private costs and benefits can a student expect? The visible costs include fees, books, transport, and lodging. The invisible costs are the opportunity costs of time, i.e. foregone earnings. The most apparent benefits of education are higher wages and lower unemployment probabilities. Non-pecuniary benefits such as better jobs, health, and personal relations may add to it as well. Because the costs of higher education are generally observable, most empirical research has focused on the benefits of higher education.

A wide range of studies estimates the private pecuniary returns to education; see Harmon et al. (2003) for a review of the evidence. Estimates are on average 8% to 10% for one extra year of education. However, these averages hide considerable heterogeneity. Estimates based on exogenous variation in schooling (e.g., based on policy reforms or date of birth) are typically higher than estimates from twin studies, which are in turn (marginally) higher than studies that rely on OLS-estimates with controls for IQ. These differences likely occur because studies that exploit variation in compulsory schooling laws estimate effects for students at the margin of high-school dropout, and extra education can strongly pay off for this particular subgroup. In addition, the estimates vary over individuals, depending on gender, IQ, birth cohort, and country. They also strongly differ across age. Returns from education are reaped from age 30 onwards and stabilize around the age of 35; see Bhuller et al. (2011) and Carneiro et al. (2003) for evidence in Norway and the Unites States. Estimates at the participation margin for higher education are typically lower than average estimates. Carneiro et al. (2011) estimate marginal treatment effects ranging from 3 to 10% for an extra year of college relative to an average return of 14%. The difference between marginal and average return could partly explain why high average returns do not necessarily lead to more participation.

There is also evidence of sizeable non-pecuniary private benefits of higher education. Education has been linked to benefits related in the realms of, for example, fertility, occupational choice, and consumption/savings behaviour; see Vila (2000); Wolfe and Haveman (2002); Johnston (2004) for overviews. Many studies look at the impact of education on health outcomes; see, for example, Grossman (2000, 2006) and Cutler and Lleras-Muney (2008) for overviews of the empirical literature. The evidence shows a strong link between education and smoking, drinking, mortality, drug use, and the prevalence of several diseases. However, the reliability of estimates differs because of the complex relationship between health and education. Causality can run both ways, and many other relevant factors are correlated to both outcomes. Conti et al. (2010) provide causal evidence for the impact of education on the prevalence of smoking. Most studies treat education as a continuous variable (years of education) and do not specifically address the role of higher education. An exception is provided by Heckman et al. (2011). They find a strong causal impact of college attendance (for both the average and the marginal individual) on several measures of health and risky behaviour, including smoking, physical health and mental health.

2.2 The second welfare theorem

The second welfare theorem assumes well-functioning markets, rational individuals, and ideal redistribution based on lump-sum transfers. Although we postpone the details of each assumption to later sections, we discuss some important consequences to understand the logic underlying the second welfare theorem.

Well-functioning markets imply that the private costs and benefits of the different individual choice options coincide with their social costs and benefits. Rationality implies that individuals choose the option with the largest private payoff, i.e. the difference between the private benefits and costs of an option. The first welfare theorem tells us that well-functioning markets and rationality together guarantee Pareto efficiency. Indeed, if each individual rationally chooses the option with the largest private payoff and if well-functioning markets guarantee that private and social payoffs coincide, then the largest social payoff results automatically (the so-called invisible hand paradigm).

Let us apply the second welfare theorem to a simple model of the participation decision in higher education. The absence of market failures implies that credit and insurance markets work perfectly. As a result, monetary costs and benefits can be shifted to other time periods and eventual risks can be perfectly insured. We can thus ignore timing and risk issues for the moment and payoffs can be interpreted as life-time payoffs with certainty.

Payoffs reflect the monetary value of both economic outcomes (like earnings and employment opportunities) and non-economic ones (such as intrinsic intellectual satisfaction, health benefits, and job satisfaction). Payoffs can be different for different types of individuals. A type will be denoted by $\theta \in \mathbb{R}$. Although the exact interpretation of a type does not matter at this stage, it can be useful to think of innate ability, acquired skills, ambitions, or socio-economic background.

Consider the laissez-faire—except for eventual lump-sum transfers—and suppose that the only decision to make is to participate or not in higher education. If someone with type $\theta \in \mathbb{R}$ chooses not to study, then a private payoff $a(\theta)$ results. If the same individual would decide to study, then an extra private benefit of higher education $b(\theta)$ is added on top of $a(\theta)$, but also the cost of higher education $c(\theta)$ has to be paid. Finally, irrespective of the decision, an individual must pay a lump-sum tax (or subsidy, if negative) denoted by $T(\theta)$. These transfers are purely redistributive, so we have $\int_{\mathbb{R}} T(\theta)f(\theta)d\theta = 0$ with $f > 0$ the density function.

An individual with type θ rationally chooses to study in the laissez-faire if it leads to a higher payoff, thus if

$$a(\theta) + b(\theta) - c(\theta) - T(\theta) \geq a(\theta) - T(\theta),$$

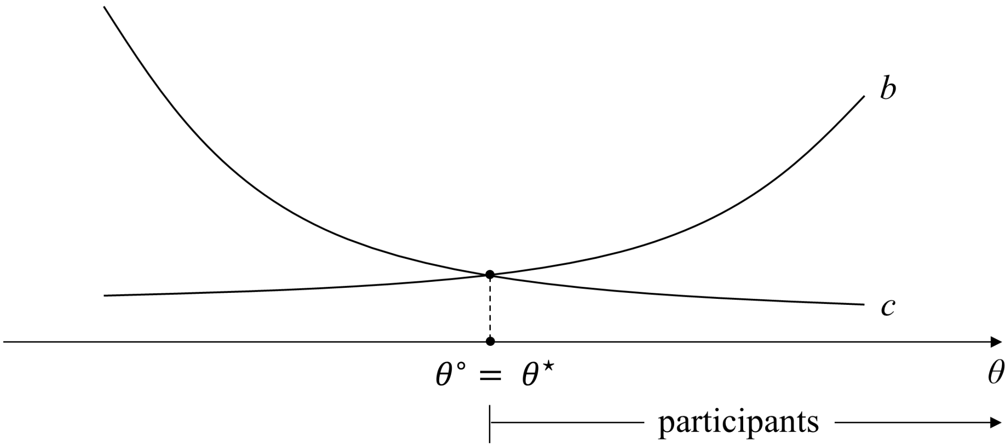
or equivalently, if

$$b(\theta) \geq c(\theta). \tag{1}$$

In other words, an individual chooses to study if the private benefit of higher education exceeds the private cost.

Figure 1 illustrates the participation decision. We assume that the functions a and

Figure 1: The participation decision



Notes: The figure shows the payoff from enrolling into higher education (b) and the cost of higher education (c), across student types θ. In this framework, all types to the right of θ° participate in higher education.

b strictly increase ($a' > 0$ and $b' > 0$) with type and that function c strictly decreases with type ($c' < 0$). Although many of the theoretical results (here and later) hold more generally, these assumptions simplify the analysis and fit rather well with the different possible interpretations of a type.[1] The simplification comes from the fact that the function $b - c$ strictly increases with type. The equilibrium set of participants is therefore fully characterized by a unique cutoff type θ° satisfying

$$b(\theta^{\circ}) - c(\theta^{\circ}) = 0. \tag{2}$$

Given that everyone maximizes his or her payoff, the equilibrium cutoff type separates students (to the right of θ°) from non-students (to the left of θ°)

Is there a reason to intervene? Although the size of the social payoff is maximal in θ°, its distribution over individuals can still be judged inequitable. The second welfare theorem states that lump-sum transfers allow to separate size and distribution. A lump-

[1]For example, a more talented individual will benefit more from participating in higher education (e.g., through complementarities between ability and effort in higher education) and will face lower study costs (e.g., because of no or less delay in study time or lower required study effort). Evidence of higher returns to education for individuals of higher (cognitive) ability is provided by Harmon et al. (2003) and specifically for college by Heckman et al. (2011).

sum transfer is a tax or subsidy of which the amount cannot be changed by changing behaviour and is therefore non-distortive. If available, then society can indeed redistribute the payoffs in any desirable way without distorting their total size. The best solution for society is therefore to maximize the average social payoff, which in the absence of market failures coincides with the average private payoff, so we must verify whether the equilibrium cutoff type maximizes the average private payoff in society.

For an arbitrary cutoff type $\theta^\star$, the average private payoff would be equal to[2]

$$\int_{<\theta^\star} (a(\theta) - T(\theta))f(\theta)d\theta + \int_{\geq\theta^\star} (a(\theta) + b(\theta) - c(\theta) - T(\theta))f(\theta)d(\theta).$$

Rearranging the average payoff as

$$\int a(\theta)f(\theta)d\theta + \int_{\geq\theta^\star} (b(\theta) - c(\theta))f(\theta)d\theta, \tag{3}$$

and differentiating it w.r.t. $\theta^\star$ leads to the first-order condition $(b(\theta^\star) - c(\theta^\star)) = 0$. In other words, the optimal set of participants is fully characterized by a unique cutoff type $\theta^\star$ satisfying

$$b(\theta^\star) - c(\theta^\star) = 0. \tag{4}$$

Compare equations (2) and (4). We immediately obtain that $\theta^\star = \theta^\circ$, i.e., the optimal set of participants and the equilibrium set coincide. So, there is no reason to intervene in higher education, at least, if the assumptions are true.

Table 1 summarizes the previous case without failures and sheds a light on the market failures and behavioural failures that we will discuss in the next two sections. The first column circumscribes each failure. The second and third column define the equilibrium and the optimality functions d° and $d^\star$ that will set the equilibrium and optimal cutoff types θ° and $\theta^\star$. The fourth column compares the resulting equilibrium and optimal participation levels, defined as P°=1-$F(\theta)^\circ$ and $P^\star$=1-$F(\theta)^\star$, with F representing the corresponding distribution function.

If there are no failures, then both functions coincide ($d^\circ = d^\star = b - c$); as a consequence, also the cutoff types ($\theta^\star = \theta^\circ$) and the participation levels ($P^\circ = P^\star$) coincide. As we will see in the next two sections, the coincidence of the functions d° and $d^\star$ disappears if we relax certain assumptions. In other words, market and behavioural

[2]In general, we must optimize over an arbitrary set of participants rather than over an arbitrary cutoff. The assumptions guarantee however that a non-connected set of participants can never be optimal so that we may safely focus on an arbitrary cutoff type (here and in the rest of the paper).

Table 1: Overview of the different failures

	$d^{\circ}(\theta) =$	$d^{\star}(\theta) =$	Part. gap
No failures	$b(\theta) - c(\theta)$	$b(\theta) - c(\theta)$	$P^{\circ} = P^{\star}$
Market failures			
- uninsurable risk			
- credit constraints			
- externalities			
- signalling			
Behavioural failures			
- misprediction			
- positional goods			
- academic peer effects			
- peer pressure			

Notes: The table shows the equilibrium (d°) and the optimality ($d^{\star}$) functions that determine the equilibrium and optimal cutoff types θ° and $\theta^{\star}$, as well as the expected participation gap (actual participation versus optimal participation). The results are (for now) only given for the benchmark case in which no failures occur.

failures will drive a wedge between equilibrium and optimal participation levels.

3 Market failures

Suppose there is no government intervention and students pay the market price of higher education up-front. The costs and benefits of higher education are uncertain and students may wish to insure against certain risks. In addition, if the market price is high, talented but poor students may require a student loan to finance their studies. Finally, the costs and benefits of higher education depend on the functioning of the education and the labour market. Each market—the insurance, credit, education, and labour market—can fail. We discuss uninsurable risks, credit constraints, externalities, and signalling.

3.1 Uninsurable risks

Students do not know the costs and the benefits of higher education for sure. On the cost side, the yearly expenses in higher education are generally observable, but whether enrolment leads to a degree and in how many years is unsure. The level of uncertainty with respect to the benefits of higher education are likely to be even stronger.[3] Income risk includes uncertainty with respect to future wages, employment opportunities, and tax legislation. The labor market consequences of graduating from college in a bad economy, for example, tend to be large, negative and persistent (Kahn, 2010). If students are risk averse and if these risks are difficult to insure, then an inefficiency arises.

Why are these risks difficult to insure? Shocks to, e.g., earnings and unemployment during an economic downturn affect the labour market as a whole. A considerable part of the income risk therefore has a collective component that is difficult to insure in a private market (Connolly and Munro, 1999). Moreover, insuring graduation risk could suffer from adverse selection and moral hazard if there is asymmetric information between insurers and prospective students about success probabilities or exerted effort in higher education.[4] Adverse selection occurs if students with a high (unobservable) success probability are not willing to insure if they must pay a premium that is based on the average success probability among students. In this case, the average success probability decreases, the premium increases, and even more students may leave to insurance scheme, and so on. Moral hazard occurs if students would exert less effort once insured against graduation risk. Adverse selection and moral hazard lead to no or incomplete insurance for a part of the student population.

If some risks cannot be insured by the market, then participation in higher education may change. The stochastic benefits and costs for an individual with type θ are now assumed to be

$$a(\theta) + \alpha,\ a(\theta) + b(\theta) + \beta, \text{ and } c(\theta) + \gamma,$$

with (α, β, γ) i.i.d. multivariate normal with zero means.[5] Individuals maximize expected utility, denoted by $Eu[\cdot]$, with a constant absolute risk aversion ρ.[6]

An individual with type θ chooses to study if the expected utility of participation is higher than the expected utility of non-participation, thus if

$$Eu\left[a(\theta) + b(\theta) + \beta - (c(\theta) + \gamma)\right] \geq Eu\left[a(\theta) + \alpha\right],$$

[3] We discuss concrete empirical evidence on misprediction of costs and benefits of higher education in section 4.1.

[4] Asymmetric information in insurance markets is first discussed by Arrow (1963) in the context of medical care.

[5] Note that β captures the total risk of $a + b$ (thus, not only of b).

[6] More precisely, $Eu[X(\theta)] = u^{-1}\left(\int u\left(x\right) f_{\theta}\left(x\right) dx\right)$ with $X(\theta)$ the stochastic payoff of type θ, f_{θ} the corresponding density function and $u : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto -\frac{1}{\rho} \exp\left(-\rho x\right)$ the VN-M utility function with $-u''/u' = \rho$ the coefficient of constant absolute risk aversion.

or, given the assumptions made, if

$$b(\theta) \geq c(\theta) + \frac{\rho}{2}\Delta\sigma, \tag{5}$$

with $\Delta\sigma$ the risk difference between participation and non-participation, i.e., $\Delta\sigma = (\sigma_{\beta\beta} + \sigma_{\gamma\gamma} - 2\sigma_{\beta\gamma}) - \sigma_{\alpha\alpha}$ with σ denoting (co)variances. In words, an individual chooses to study if the private benefit of higher education is sufficient to cover the private cost augmented by a risk premium. The risk premium is equal to half the risk-aversion coefficient ($\rho/2$) multiplied by the risk difference $\Delta\sigma$. If we compare equation (1) with (5), then the curve $b - c$ in Figure 1 will shift up or down depending on the sign of the risk premium.

We assume for the moment that individuals are risk averse. The sign of the risk difference $\Delta\sigma$ will then tell us whether uncertainty causes more or less participation. It is worth stressing that, irrespective of the impact of uninsurable risk on participation in higher education, there is always an inefficiency. Even if $\Delta\sigma$ is equal to zero, while there is clearly no impact on participation, participants and non-participants would both still benefit from insurance.

It is a priori not obvious whether the risk difference is positive or negative. On the benefit side, for example, participation in higher education may lead to a higher wage volatility, but at the same time it may also reduce the risk of unemployment. Meghir and Pistaferri (2004) state that "the higher returns from increased education come at the cost of higher earnings risk." This suggests that $\sigma_{\beta\beta} > \sigma_{\alpha\alpha}$. If the covariance $\sigma_{\beta\gamma}$ is not too high—a priori we expect it to be negative—then the risk difference $\Delta\sigma$ will be positive. Risk then leads to lower participation. Carneiro et al. (2003) confirm that less uncertainty leads to more participation. They show that a large part of the heterogeneity in the rate of return is not predictable at the time of the participation decision.[7] Eliminating uncertainty increases participation. If individuals were to know all the costs and benefits (including non-pecuniary ones) of higher education for sure, then 12 percent of those with only high school education would attend college, and 2 percent of college students would not have enrolled.

Up to now, we assumed that individuals are risk averse. Belzil and Leonardi (2007) study the relation between risk aversion and schooling decisions and identify, by using a lottery experiment, that about 95% of their respondents, comprised of a sample of

[7] The authors estimate that knowing ones cognitive type predicts around 30% of the monetary return of college over high school.

Italian households, is risk averse. However, they also point to large differences between individuals. Belzil and Hansen (2004) directly infer risk preferences from schooling decisions and identify a modest degree of risk aversion. In theory, if the risk difference $\Delta\sigma$ is positive, then individuals with a higher risk aversion should be less likely to participate in higher education. Belzil and Leonardi (2007) confirm this hypothesis and show moreover that the explanatory power of risk aversion is substantial, i.e., similar to that of parental education. Chen et al. (2001) also identifies negative effects of risk aversion on participation in the United States, but these are smaller in magnitude.[8]

Insurance can also induce moral hazard, e.g., studying or working less hard if the insurer bears the risk. Although there is no direct evidence that insurance provision leads to moral hazard, there is evidence that study effort and performance vary with changes in costs or benefits.[9] Some studies analyze how student effort reacts to changes in costs. Garibaldi et al. (2012) demonstrate that, in Italy, a 1000 euro increase in tuition for delayed years reduces the probability of delayed graduation by 5.2 percentage points, without changing student performance. Fricke (2013) estimates that an announced increase in tuition for delayed years at a Swiss university, roughly equivalent to 800$, lead to a 9.7 percentage point increase in graduation for students near the end of their studies, while their grades remain similar. Other studies estimate the effects of directly incentivizing students with financial rewards (which effectively changes the benefit side of the participation decision). The results are rather mixed. In a study conducted at the University of Amsterdam, Leuven et al. (2010) identify that incentives improve the performance of high-ability students, but reduce low-ability students' performance.[10] Although the reward was only tied to first-year academic performance, treatment effects even persisted into later years (including increased dropout for low-ability students). Angrist et al. (2009) find that financial rewards for high performance bring positive and persistent results, but only when the rewards were combined with the provision of academic support services and only for women. Also, students made much stronger use of support services when they also could receive financial rewards for high performance.

[8] These findings are not causal, however, and prone to reverse causality.

[9] We focus on study performance; moral hazard in the labour market will be discussed in Section 5. Note also that there is no hard evidence for adverse selection.

[10] The authors suggest that low-ability students are demotivated because the common achievement threshold is not attainable for them.

3.2 Credit constraints

Income risk can in turn lead to default risk, i.e., the risk of not being able to repay a loan that is used to finance higher education. Evidence on default rates shows that they are substantial and differ by type of institution. Steinerman et al. (2011) report that the three-year default rate for students that enrolled in 2008 in the United States is equal to 7.6 percent for private non-profit institutions, 10.8 percent for public institutions, and 24.9 percent for private for-profit institutions.

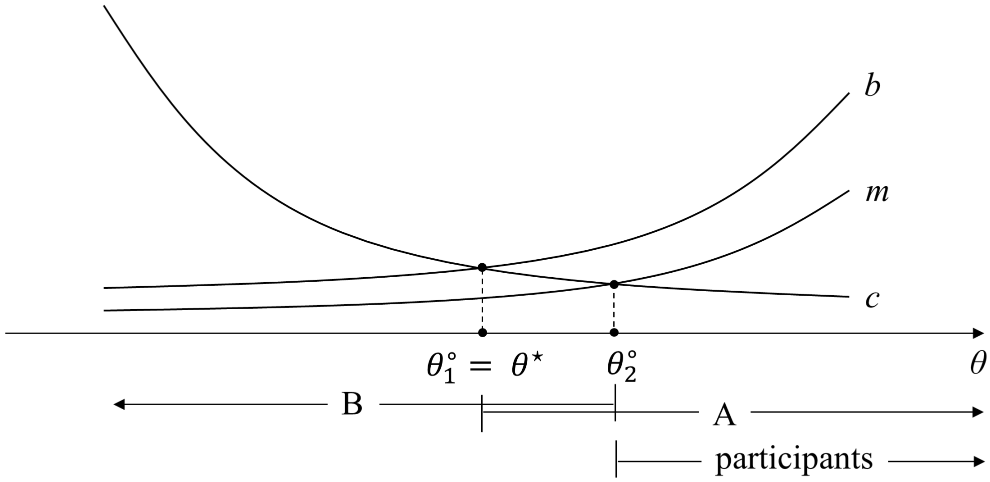
Like income risk, default risk is difficult to insure in a private market. As before, hidden asymmetric information about the default risk of students may cause adverse selection (low risk students are driven out of the market by high risk students) and moral hazard (students exert less work effort to pay back the loan). In addition, students have little collateral to offer in a non-slave state to secure their loan (Friedman, 1955) and parents are often reluctant to provide security (Mazzeo, 2007). The lack of insurance may imply that capital markets do not provide sufficient credit. Especially poor but otherwise talented students may be financially constrained and refrain from participation (Gross et al., 2009).

We illustrate how moral hazard can lead to 'rational' default on loans, which, if anticipated by lenders, may result in credit constraints. Suppose a student can borrow money in a private credit market to finance his study cost $c(\theta)$.[11] As before, it is rational for an individual with type θ to study if $b(\theta) \geq c(\theta)$, i.e., if the benefit of higher education is sufficient to cover the costs. Define θ_1° the cutoff type who is indifferent between participating or not, so θ_1° must satisfy $b(\theta_1^\circ) - c(\theta_1^\circ) = 0$.

We follow Lochner and Monge-Naranjo (2011) and use a short-cut to introduce moral hazard in the model. Suppose there is a default cost, denoted by $m(\theta)$ for an individual with type θ. The default cost increases with type, and thus, also with earnings. One motivation for this assumption is that default restricts access to further credit, which is more harmful to richer people. Another is that subsistence incomes are protected from the lender in many countries. Leaving the labor market leads to an income loss– the difference between your market income and the subsistence income–but also allows you not to pay back your loan. The income loss can therefore be considered as the default cost, which indeed increases with type.

[11] Study costs refer here to the costs net of financial support from parents. It is natural to assume that support increases with type, which strengthens the decreasing nature of the cost function by type.

Figure 2: Credit constraints



Notes: The figure denotes the payoff (b) and the costs (c) of higher education, as well as the cost of default (m), across student type θ. Zone A indicates students for whom it is rational to participate using a loan, while zone B indicates students for whom it is rational to default if they take a loan. For types in the intersection of both zones (between θ_1° and θ_2°), it is rational to participate using a loan and subsequently default. If lenders anticipate this and constrain credit, participation will reduce to those to the right of θ_2°.

Default is rational if the default cost is lower than the loan repayment, i.e., if

$$m(\theta) < c(\theta),$$

We define θ_2° as the cutoff type that is indifferent between default and repayment, satisfying $m(\theta_2^\circ) = c(\theta_2^\circ)$.

If $\theta_2^\circ \leq \theta_1^\circ$, then those who want to participate (to the right of θ_1°) will never default on their loan (i.e., are never to the left of θ_2°). Hence, no problem occurs. Figure 2 depicts the alternative case where $\theta_2^\circ > \theta_1^\circ$. In this situation, some individuals with a type above θ_1° require a loan to study, but only for those above θ_2° it is rational to pay it back. Lenders would make a loss.

If lenders anticipate rational default and if type is private information, then lenders could constrain credit to $\bar{c} = c(\theta_2^\circ)$ to avoid losses.[12] Only individuals with a type above

[12]Lenders could also increase the interest rate by adding a risk premium. This will increase the total costs of higher education and therefore participation will decrease. Some of those that previously defaulted on their loans will now not participate, but additional defaulters are created as well since the default cost c increases. If the latter outweighs the former, it will further increase interest rates, reduce participation and increase default. If the former outweighs the latter, default is reduced, but at the expense of lower participation in higher education.

θ_2° can gather sufficient resources to study and lenders do not incur losses anymore. Moral hazard leads to an inefficiency: the benefits of higher education are larger than the costs for individuals with types between θ_1° and θ_2° , but the credit constraint prevents them to study.

There is ample empirical evidence on credit constraints. The literature distinguishes between short-run and long-run constraints. The former refers to a lack of means to directly finance the costs of higher education. The latter refers to a lack of means to make investments earlier in life that improve the learning process of the child and therefore improve his or her future access to and return from participation in higher education. Research finds strong and increasing correlations between family income and college attendance in the United States. Carneiro and Heckman (2002) show that the difference in college attendance between the top half and the bottom quantile of parental family income is around 20% in the 1980s and 30% in the 1990s. But this does not tell us which type of constraint is mainly responsible for the observed attendance gap. The available research generally indicates that long-run constraints are dominating. Cameron and Heckman (2001) show that the effect of income on college attendance is low once we correct for maternal ability. If maternal ability is a good proxy for long-run family factors like genetics, early family environment and previous school quality, then long-run constraints are crucial. Carneiro and Heckman (2002) also compare such conditional enrolment rates and conclude that at most 8% of the population faces short-run credit constraints with respect to college enrolment. Duncan and Brooks-Gunn (1997) confirm that long-run constraints are far more important than short-run credit constraints for college attendance. The children of wealthy and highly educated parents perform better (by nature and nurture) at compulsory school, which is a crucial proviso for college access and success.[13] Keane and Wolpin (2001), in contrast, use a structural model and show that short-term credit constraints do exist and are tight (in the sense that students cannot finance college costs using uncollateralized loans). Yet, in line with the previous evidence these constraints turn out to have little effect on college attendance. Their simulations show that relaxing the constraints mainly affects other choice margins, e.g., students work less while in college and consume more.

More recent findings suggest however that the estimates of short-run credit constraints are sensitive for several reasons. First, there are indications that constraints have

[13]Early childhood interventions are effective and efficient to address such long-run constraints; see, e.g., Cunha et al. (2006).

increased over time. Belley and Lochner (2007) use more recent data and show that the effect of parental income on participation (conditional on ability) is twice as high in the early 2000's compared to the 1980's in the United States. Winter (2014) confirms that the enrolment gap between children from rich and poor families is widening over time. This is not caused by tighter borrowing constraints, but rather higher earnings risk over time which has made especially low- and middle-income parents less inclined to fund the college participation of their children.

Second, the estimates are very sensitive to the definition of family income. Lochner and Monge-Naranjo (2011) indicate that the combined effect of income and wealth is twice as large as the effect of income alone. External financial aid also plays a role. Winter (2014) argues that conditional enrolment rate differences underestimate short-run constraints. These conditional rates should be *higher* for children from poor families in the absence of credit constraints because they receive more financial aid. Additionally, Hilger (2014) shows that negative parental income shocks just before enrolment have only a minor impact on participation, as children from poor families receive little college contributions from their parents anyway and what they lose is offset by increased financial aid.

Third, credit constraints are different in Europe. Chowdry et al. (2013) find a substantial enrolment gap between children from rich and poor families in England, but the gap almost completely disappears after controlling for secondary school achievement (conditional gaps are 1.0 percentage points for males and 2.1 percentage points for females). This is not surprising as the private costs of higher education are much lower in Europe.[14] However, if European policies shift further in the direction of increased tuition, then the estimates of credit constraints could move towards those found in United States. As we will see in Section 6.2, this also depends on whether tuition increases are supplemented with additional grants and loans.

3.3 Externalities

Up to now, we assumed that the decision to participate in higher education affects only individual payoffs. This assumption is compromised if the participation decision directly affects the utility of other individuals in society. If there exist so-called externalities, then the private and social payoff of participation differs, leading to an ineffi-

[14]The cohort studied in Chowdry et al. (2013) was eligible for higher education just before the major reforms and increased tuition rates that were introduced in the United Kingdom from 2006 onwards.

ciency.[15]

Positive externalities in higher education could occur if graduation increases the productivity of all individuals in society (Lucas, 1988). Graduates can also accelerate economic growth via learning-by-doing, technological diffusion, and innovation (Arrow, 1962; Nelson and Phelps, 1966; Romer, 1993). Also non-market externalities may exist for society in terms of, e.g., crime, public health, the environment, and social cohesion. If positive externalities exist, then there will be too little participation in higher education from a societal point of view.

To see this, suppose an individual with type θ has an external effect $e(\theta)$ on everyone else in society in case he would study and graduate. The total externality is measured by $E = \int_{\geq\theta^\circ} e(\theta)f(\theta)d\theta$, with θ° the equilibrium cutoff type. If we add E on both sides of equation (1), then the participation decision remains unchanged and the equilibrium cutoff is again defined by equation (2), i.e., $b(\theta^\circ) - c(\theta^\circ) = 0$.

The equilibrium participation level will not be efficient however. For an arbitrary cutoff type $\theta^\star$ the average social outcome is equal to

$$\int_{<\theta^\star} (a(\theta) + E - T(\theta))f(\theta)d\theta + \int_{\geq\theta^\star} (a(\theta) + b(\theta) + E - c(\theta) - T(\theta))f(\theta)d\theta.$$

Using the definition of E, we can rearrange it as

$$\int a(\theta)f(\theta)d\theta + \int_{\geq\theta^\star} (b(\theta) + e(\theta) - c(\theta))f(\theta)d\theta.$$

Differentation leads to the first-order condition $b(\theta^\star) + e(\theta^\star) - c(\theta^\star) = 0$. In words, an individual with type θ should study from a social point of view if the social benefit of higher education—the private benefit $b(\theta)$ augmented by the external effect $e(\theta)$—exceeds the private cost $c(\theta)$. If external effects are positive, then indeed too little participation results. Coming back to Figure 1, the external benefit *e* would be added to the private benefit *b*, creating a (social) benefit function *b*+*e* that would lie above *b*. The efficient optimum $\theta^\star$ would then lie to the left of the equilibrium θ°; because the individual does not take his or her external effect on others into account when making a participation decision, too little participation results.

Macro-economic studies claim that the average educational attainment of a country has positive effects on macro-economic performance over and above the enhancement of

[15] Peer interactions are similar to externalities and will be introduced in Section 4.

individual productivity. Sianesi and Reenen (2003) review the evidence and conclude that there is some justification for this hypothesis, but that the estimates are likely to grossly overstate the true effects. Krueger and Lindahl (2001) argue that estimates of the effect of education on growth rely on restrictions that are rejected by the data. Their results especially cast doubt on the added value of higher education. This view is largely shared by Lange and Topel (2006), who state that the evidence on large social returns of education is mixed, and that the private benefits outweigh the spillovers by far. Moretti (2004) takes a different approach by looking at productivities at the plant level. He finds that less educated workers are more productive in locations with more highly educated workers. There is little evidence whether pecuniary spillover effects differ by type of post-secondary education. Murphy et al. (1991) identify a positive correlation between growth and the share of engineering majors in a country, while the correlation with the share of legal majors is negative. More precisely, an increase in the share of engineering majors of 10% (which would double the current share) would have increased the 1970-1985 growth rate by 0.5%. However, these results are based on small samples and border on statistical significance.

There is also empirical evidence linking education to non-pecuniary spillovers on public health, crime, pollution, and social cohesion; see McMahon (2004) for an overview of macro-economic evidence. Recent studies have shown that those links are often causal. Heckman et al. (2011) demonstrate that attending college leads to a higher probability of voting and higher trust. Other studies also find an impact of education on voter turn-out, but these effects tend to be lower in Europe than in the United States (Milligan et al. (2004), Dee (2004), Siedler (2010)). Lochner and Moretti (2004) and Machin et al. (2011) show that compulsory schooling law changes lead to lower crime rates in both the United States and the United Kingdom. Furthermore, spillover effects can also be intergenerational. Currie and Moretti (2003) use variation in college proximity as an instrument and find that parental education also has an intergenerational effect on infant health. Holmlund et al. (2011) use twin and adoptee data as well as policy reforms in Scandinavia and the United States to show that more educated parents 'produce' more educated children. The causal estimates indicate an increase of child's educational attainment of 0.1 years for every extra year of schooling for the parent. Recent evidence from Sweden has identified intergenerational effects in educational attainment across four generations (Lindahl et al., 2012). This study finds that the effect of the educational attainment of grandparents and great-grandparents is two to three times larger than the effect predicted by the correlations from one generation to the next.[16] In other words, grandparents have an independent transmission effect on the completed years of education of their grandchildren.

3.4 Signalling

The view that higher education can signal productivity, besides enhancing it, has been raised first by Arrow (1973) and Spence (1973). We distinguish between the strong and the weak signalling hypothesis depending on whether higher education only signals productivity (strong) or whether it both signals and enhances productivity (weak).[17] To explain the inefficiency, let us focus for the moment on the strong hypothesis. Suppose that employers do not observe individual productivities, a case of asymmetric information. Signalling occurs if high productivity individuals can communicate their unobserved productivity to employers in a credible way, i.e., such that low productivity types will not mimic their behaviour. Costly investment in higher education offers a plausible signal. Yet, signalling only transfers income from low to high productivity individuals and the signal itself is costly without enhancing productivity (in the strong hypothesis). Signalling leads therefore to too much participation from a societal point of view.

To show the result more generally, suppose from now on that the weak signalling hypothesis holds: higher education enhances earnings as before, but it can also be used to credibly signal unobservable ability. The average social outcome is equal to equation (3). So, from a societal point of view, the efficient cutoff $\theta^\star$ uniquely satisfies $b(\theta^\star) - c(\theta^\star) = 0$.

The presence of asymmetric information in the labour market will lead to inefficiencies. If employers cannot observe individual types (productivities), they can only differentiate wages between graduates and non-graduates. Suppose that employers correctly anticipate the average productivity of non-graduates and graduates. Rational expectations in a competitive labour market imply that the difference in wages will reflect the difference in average productivities, being

$$\Delta w(\theta^\circ) = E[a(\theta) + b(\theta)|\theta \geq \theta^\circ] - E[a(\theta)|\theta < \theta^\circ], \tag{6}$$

[16] For example, the direct transmission coefficient from great-grandparents equals 0.145, while multiplying the three transmission coefficients from one generation to the next gives a coefficient of 0.05.

[17] Higher education may increase productivity by direct learning, but also by sorting individuals (Stiglitz, 1975). For example, if one individual has a comparative advantage in engineering and another in brain surgery, then society gains by putting the right person in the right place.

given the equilibrium cutoff type θ° and $E[\cdot]$ the (conditional) expectation.[18] For clarity, note that the function Δw is a function of the equilibrium cutoff type and is thus the same for all types. For later use, also note that the function Δw is not necessarily strictly increasing, but must lie everywhere strictly above the function b.[19]

An individual with type θ that correctly anticipates the level of participation will study if the resulting wage difference $\Delta w(\theta^\circ)$ exceeds the cost of higher education $c(\theta)$. Because Δw is not necessarily strictly increasing in θ°, several equilibria may occur. Each equilibrium cutoff must satisfy

$$\Delta w(\theta^\circ) = c(\theta^\circ). \tag{7}$$

Take an arbitrary equilibrium cutoff θ°. Because we have seen before that the wage difference $\Delta w(\theta^\circ)$ is larger than the benefit $b(\theta^\circ)$, equation (7) implies $b(\theta^\circ) - c(\theta^\circ) < 0$. As the efficient cutoff $\theta^\star$ solves $b(\theta^\star) - c(\theta^\star) = 0$, we have $b(\theta^\circ) - c(\theta^\circ) < b(\theta^\star) - c(\theta^\star)$ and each equilibrium cutoff θ° must lie below the efficient cutoff $\theta^\star$. Hence, signalling always leads to too much participation from a societal point of view.

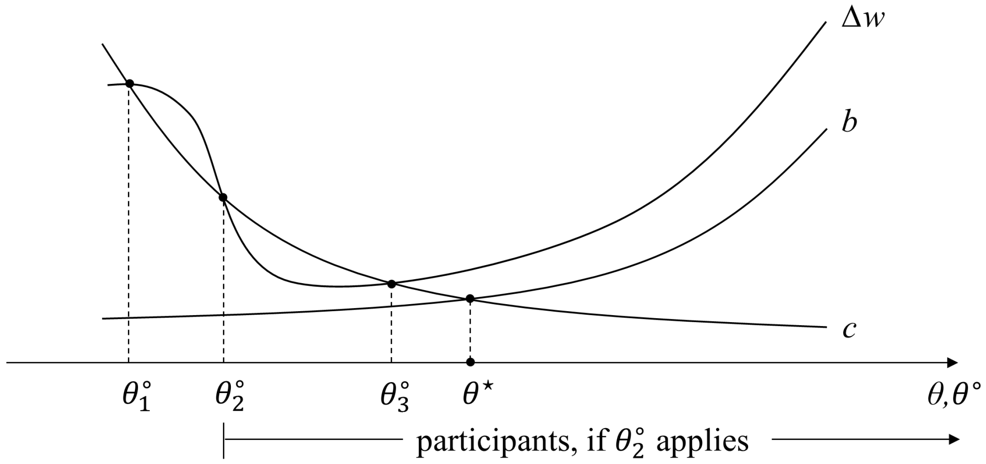
Figure 3 illustrates the function Δw. Because Δw is not strictly increasing, multiple equilibria exist to the left of the optimal cutoff in this specific example.

Until now, we analyzed participation in higher education as a signal for ability and ambition. Studying longer can be interpreted in a similar way. Master and advanced master programs can be used to differentiate oneself from bachelor and master students. Loosely speaking, the longer one studies, the more signalling inefficiencies are cumulated, and the more reason there is to intervene. This could be an argument to differentiate tuition between bachelors, masters, and advanced masters programs. The same is true for similar courses at different universities. Course contents can be exactly the same, but fee differences imply that the more expensive program can be used to signal ability/ambition. Thus, signalling can give market power to universities allowing them to charge higher fees for the same quality of education. To put it differently, Bertrand competition between universities may raise tuition if signalling is part of the game.

[18] More precisely, $E[z(\theta)|\theta \in S] = \int_S z(\theta) f(\theta) d\theta / \int_S f(\theta) d\theta$ for some set $S \subseteq \mathbb{R}$.

[19] $\Delta w > b$ follows from combining $E[a(\theta) + b(\theta)|\theta \geq \theta^\circ] > a(\theta^\circ) + b(\theta^\circ)$ and $E[a(\theta)|\theta < \theta^\circ] < a(\theta^\circ)$.

Figure 3: Signalling



Notes: The figure denotes the benefits (b) and costs (c) of education, across student type θ, as well as the wage differential from attending higher education (Δw), across the cutoff type θ°.

It is unlikely that earnings payoffs through education are all about signalling.[20] But the rejection of the strong signalling hypothesis does not tell us that signalling cannot occur. The early literature has used several strategies to test for the presence of signalling. For example, Layard and Psacharopoulos (1974) started a long-running debate on whether obtaining a degree (a 'sheepskin') has an effect on earnings over and above the effect of the number of years in education. This would be the case if credentials signal desirable traits like perseverance to employers. Overall, the early and recent evidence turns out to be mixed and varies over estimation strategies and countries; see Sessions and Brown (2005) for a comprehensive overview. Sheepskin effects have been found in the United States (Hungerford and Solon, 1987; Belman and Heywood, 1991; Jaeger and Page, 1996) and New Zealand (Gibson, 2000). However, multiple other studies contradict these conclusions. Patrinos (1996) and Griffin and Cox Edwards (1993) do not find evidence of sheepskin effects for Guatemala and Brazil, respectively.

Natural experiments are better equipped to obtain valid estimates of the presence of signalling in the presence of endogenous schooling decisions. Lang and Kropp (1986)

[20]Direct causal evidence of a productivity effect of education is scarce as it is difficult to disentangle signalling and productivity empirically. A large bulk of indirect evidence exists. For example, Oosterbeek (1992); Groot and Oosterbeek (1992) show that extra years of education caused by delay or without achieving a degree still lead to significant returns in the labour market. This evidence is in line with productivity theory and opposed to signalling.

and Chevalier et al. (2004) use changes in compulsory schooling laws to test for the presence of signalling in the United States and the United Kingdom. If signalling were prevalent, one would expect these reforms to increase educational attainment throughout the distribution. Low-ability individuals experience a compulsory increase in their completed years of schooling, requiring high-ability individuals to distinguish themselves again. The two studies obtain opposite results: the evidence in the United States is consistent with the signalling hypothesis (enrolment rates for unaffected age groups increased), while the signalling hypothesis is rejected for the United Kingdom. One possible reason is that the tuition differences in the United Kingdom are smaller and therefore less prone to signalling compared to the United States. Finally, Bedard (2001) finds that local labour markets without university have lower dropout rates at high school. This is consistent with signalling, but not with human capital theory. If there is no university, then low ability students have an extra incentive to complete high school because completion makes their educational credentials indistinguishable from high-ability students who are constrained in entering college.

4 Behavioural failures

Up to now we assumed that individuals behave like a homo economicus, a decision-maker that rationally weighs the costs and benefits of each action before choosing the best one. Research in behavioral economics analyzes how individuals make decisions in real-world settings and has flourished over the past decades; see, e.g., Rabin (1998, 2002) and DellaVigna (2009) for and overview. It shows that the homo psychologicus, with bounded rationality, bounded will-power, and subject to social interactions, often provides a better description of decision-making. The participation decision made by students and their parents is probably no exception.

First, individuals may mispredict the costs, benefits, and probabilities that are relevant to the participation decision. Besides providing empirical facts about misprediction, we also discuss plausible mechanisms such as a simple lack of information, complexity of application, and psychological mechanisms. Second, individuals do not make decisions in isolation. Individual decisions influence the decisions of others, for example through social status and conformity. Such effects are labeled as 'social interactions'. Both mispredictions and social interactions can lead to inefficiencies as the actual decisions do not necessarily promote the societies' best interests.

Before proceeding we stress that caution is needed for at least three reasons (Camerer et al., 2003). First, behavioural economics is still in its infancy, especially in higher education. A rare overview of behavioral economics in an educational context is provided by Jabbar (2011). Some experimental results need further exploration and experimentation; and some of the behavioural mechanisms that can explain these facts are speculative at this stage. Second, it is not always clear whether behavioural failures are truly failures. Individuals often identify themselves with 'wrong' choices. Therefore, government intervention must often balance between libertarianism (respecting 'wrong' choices if individuals identify themselves with these choices) and paternalism (correcting 'wrong' choices, even if individuals identify with these choices). Third, behavioural mistakes are far from universal. Policies should try to remove irrational choices, while safeguarding as much as possible the choices made by rational individuals.

4.1 Misprediction

A rational participation decision requires knowledge of expected costs and benefits. But the evidence shows that mispredictions occur. We first show the evidence and discuss the consequences for participation behaviour. We then discuss several mechanisms that could possibly explain misprediction.

4.1.1 Facts about misprediction

Research shows that students tend to overestimate the costs of higher education. Students and parents in the United States have an upwardly biased estimate of college costs, especially for public education (Horn et al., 2003). Grodsky and Jones (2007) identify an average overestimation of 75% of the true yearly tuition fee.

Benefits, on the other hand, are more likely to be underestimated. A study on Canadian students shows that they underestimate the annual income differential between high school and university graduates (Usher, 2005). Students are also largely unaware of financial aid, and this applies even stronger to those who are eligible (Chan and Cochrane, 2008). They often assume incorrectly that their family income is too high or that good grades are required to be eligible (Matus-Grossman and Gooden, 2002; Zarate and Pachon, 2006).

Suppose that $\hat{b}(\theta)$ and $\hat{c}(\theta)$ are predictions of benefits and costs of higher education. An individual with type θ will study if the predicted benefit $\hat{b}(\theta)$ outweighs the predicted cost $\hat{c}(\theta)$; while he should study, from an efficiency point of view, if the true benefit $b(\theta)$ is larger than the true cost $c(\theta)$. The evidence tells us that the benefits of higher education are underestimated and the costs are overestimated; i.e., $\hat{b} - \hat{c} < b - c$. In relation to Figure 1, this means that the predicted net benefit curve lies below the true one and therefore too little participation will result (since, $\theta^{\circ} > \theta^{\star}$).

Misprediction of costs and benefits could be associated with parental background. Horn et al. (2003), for example, report that the probability that parents accurately estimate tuition fees is associated with information exchange between parents. If social networks are stronger for socio-economically advantaged groups, more and better information might lead to a more reliable prediction for those families. The association between misprediction and parental background appears complicated. Grodsky and Jones (2007) show that ethnic minorities and parents with low income overestimate yearly tuition fees with the same percentage as the rest of the population, but they are less likely to provide an estimate of college costs altogether. Moreover, the variance of their estimates is significantly larger. Hence, there is some evidence that misprediction disproportionably affects those from weaker family backgrounds.

4.1.2 Plausible mechanisms

To design appropriate policies to counteract misprediction, it is crucial to understand where it comes from. Is it a lack of information? Is complexity of application procedures, combined with bounded rationality, part of the problem? Or are there deeper psychological mechanisms at play?

If misprediction is simply a matter of information, then the appropriate policy would follow immediately: provide the correct information. However, the available evidence suggests that providing students and their families with more accurate information does not necessarily solve the problem. Avery (2010) shows that providing counseling in the application process has little or no effect on the quality of the applications and that a large share of students did not follow up on the counselor's advice on colleges that appeared the best matches. Bettinger et al. (2012) show that providing students with help in the application for students aid is only effective when the provided information is combined with help in completing the application. An 'information only' treatment had no effect on the number of applications, college attendance, and persistence.

Even if all information is available, bounded rationality may also cause too little participation. In the context of welfare benefits, Kleven and Kopczuk (2008) show that program complexity can explain incomplete take-up. As students often report that the complexity of students aid programs discourages them from applying, it is likely that the same mechanism applies. Dynarski and Scott-Clayton (2006) argue that reducing complexity in the application process for student aid would lead to a more efficient and more equitable allocation. Another instance of bounded rationality is related to college application. Pallais (2015) exploits a reduction in the costs of sending score report cards for the ACT, a college entrance exam, to prospective colleges to assess student applications. A very small cost reduction (students got the fourth score report for free, while it used to cost 6$) had a strong impact on the number of applications sent and is likely to increase the chance of entering (selective) colleges. This behavior is inconsistent with traditional economic theory, as the returns of sending the extra application, which greatly increases the chance of attending selective colleges, are large. Providing students with three free score reports in the initial situation likely 'primed' them to send exactly three applications.

Finally, there could also be psychological reasons underlying misprediction. We discuss three deviations from standard economic theory as described by DellaVigna (2009): non-standard (time and risk) preferences, non-standard beliefs (overconfidence), and non-standard decision-making (framing).

Non-standard time preferences, e.g., a preference for immediate over delayed utility, is one possible explanation for misprediction; see Harrison et al. (2002) for evidence. Higher education is an investment with immediate monetary and non-monetary costs, and delayed benefits. Overweighting the current costs and underweighting the future benefits can reduce participation in higher education. Individuals then fail to do in the short term what is optimal for them in the long-run.

Non-standard risk preferences can also play a role. Kahneman and Tversky (1979) argue that individuals tend to be risk averse for gains w.r.t. a reference situation and risk loving for losses. Page (2005) proposes this argument in the context of education, where aspiration levels could operate as reference points. If the education level of one's parents is the relevant reference, then education is a gain for individuals with low educated parents, and a loss for individuals with high educated parents, resulting in too low participation for the former group and too high participation for the latter group. Page et al. (2007) show experimental evidence that framing of reference points as either gains or losses affects choices, although this study does not show whether aspiration levels indeed serve as reference points in an educational setting.

Additionally, non-standard beliefs, e.g., overconfidence, may occur. Students, especially low performers, tend to overestimate their skills. Grimes (2002) and Nowell and Alston (2007) find both that students with lower GPA's exhibit a higher degree of overestimation of their actual performance. If overconfidence implies that low performing students also overestimate their probability of success in higher education, then overconfidence goes against the evidence on misprediction: it should lead to an overestimation of expected benefits and an underestimation of expected costs.

Finally, framing can provide an explanation for some of the mispredictions. The presentation of the choice options can matter, even if the choices are otherwise equivalent. For example, tuition and grants are paid and received at approximately the same time. A combination of high tuition and high grants can be made financially equivalent to a combination of low tuition and low grants. But if tuition (the sticker price) is a more salient feature, then too much weight is put on tuition relative to grants, and the second combination leads to more participation.

The evidence on framing effects in higher education is mixed. Heller (1997) finds some evidence that students in the United States have distinctive responses towards changes in tuition, grants, loans, and work studies. More specifically, enrolment is more sensitive to changes in tuition than to changes in financial aid. Within the latter group, changes in grants have stronger effects than changes in loans or work studies. These results are not fully consistent across studies and opposite findings are relatively frequent. Falch and Oosterbeek (2011) review the evidence in Europe and conclude that responses towards aid and tuitions have similar elasticities. One plausible reason is that aid is more universal and less complex in Europe compared to the United States. In addition, there is substantial heterogeneity in sensitivity to framing. Heller (1997) identifies higher sensitivities for low-income students and students in community colleges. In addition, framing sensitivity is also higher for individuals with lower cognitive ability; see, Peters et al. (2006). This is suggestive evidence that framing could be (partly) the result of bounded rationality. Still, it is not exclusive to those with relatively lower cognitive ability. Avery and Hoxby (2004) find similar 'irrational' responses to framing for a selected group of high-aptitude students. Only those high-aptitude students with parents that had high income and attended selective colleges in the past behaved in a way that is consistent with economic theory.

Other framing effects like loan aversion imply that credit take-up by students is lower if credit is explicitly labelled as a loan, ceteris paribus. Caetano et al. (2011)

finds, in an experimental setting among Latin-American students, that individuals shy away from offered contracts that involve loans. They also show that this effect mainly operates through labelling; participants behave more rational when the choice is framed as a 'human capital contract' that is financially equivalent. Johnson and Montmarquette (2011), on the other hand, find little evidence of debt aversion in a study among Canadian students.

4.2 Social interactions

In Section 2 we assumed that the participation decision of an individual does not affect other individuals in society. Externalities, discussed in Section 3, were a first counterexample. Note however that externalities directly influence the payoff of other individuals, but do not influence their participation decisions. In other words, the externalities as defined in Section 3 do not lead to interactions. Social interactions occur if the participation decisions of some individuals influence the decisions of other individuals. As a consequence, they may also influence the well-being of others and inefficiencies may result. We discuss interactions caused by positional goods and peer interactions. Peer interactions can follow from interactions in production (academic peer effects) and preferences (peer pressure).

4.2.1 Positional goods

Suppose that individuals derive utility from social status. Social status is typically defined on the basis of the position that individuals occupy in society. The proportion of non-graduates and graduates contains information about your position. In case you graduate, the proportion of non-graduates indicates the (minimal) number of individuals with an inferior type. Similarly, in case you do not graduate, the proportion of graduates can be interpreted as the (minimal) number of individuals with a superior type. We show that if individuals strive for social status, then too much participation results.

Suppose individuals receive a social status gain upon graduation, i.e., an extra benefit that linearly increases with the proportion of non-graduates in society. Recall that P°=1-$F(\theta)^{\circ}$. The social status gain is then equal to $s(1-P^{\circ})$, with $s>0$ measuring the importance of social status. In the same vein, suppose that non-graduates experience a status loss that increases with the proportion of graduates, i.e., the loss is equal to $-sP^{\circ}$. If individuals correctly anticipate the proportion of graduates, then an individual with type θ will study if

$$a(\theta) + b(\theta) + s(1 - P^{\circ}) - c(\theta) \geq a(\theta) - sP^{\circ},$$

or equivalently, if

$$b(\theta) + s \geq c(\theta).$$

The resulting cutoff type θ° is defined by $b(\theta^{\circ}) + s - c(\theta^{\circ}) = 0$.

But this result cannot be optimal. Note that social status is (deliberately) modelled as a zero-sum game: the total payoff for graduates $P^{\circ} \times s(1 - P^{\circ})$ is exactly offset by the total loss to non-graduates $(1 - P^{\circ}) \times (-sP^{\circ})$. The average outcome in society is thus the same as in equation (3) and the efficient cutoff $\theta^{\star}$ must satisfy equation (4). Given $s > 0$, we must have $\theta^{\circ} < \theta^{\star}$ and social status leads indeed to too much participation. Applied to Figure 1, *b* now represents the social benefit curve which lies below the private benefit curve faced by the individual (*b*+*s*). The resulting equilibrium will therefore lie to the left of $\theta^{\star}$.

There is empirical evidence that people value the social status effect of higher education. Solnick and Hemenway (2005) and Solnick et al. (2007) analyze the extent to which different goods are positional, i.e., whether the utility derived from specific goods depends on the consumption level of others. Their survey data show that education also exhibits a relatively high degree of positionality. In addition, there is abundant evidence in the happiness literature that income is a positional good. Happiness increases with personal income, but decreases with reference income such as past personal income or mean income in society (Clark et al., 2008). Because higher income can be obtained via higher education, the positionality of income may indirectly lead to too much participation in higher education.

Finally, if jobs depend positionally on education levels, then lower skilled workers would be replaced by higher skilled workers during recessions (the so-called crowding-out hypothesis). The evidence is mixed. For the Netherlands, Teulings and Koopmanschap (1989) support the crowding-out hypothesis, while Gautier et al. (2002) reject it. Pollmann-Schult (2005) finds evidence in West Germany that the competition between low and high skilled workers for low-skilled jobs increases during recessions. A recent study by Beaudry et al. (2013) indicates that high-skilled workers have pushed low-skilled workers down the occupational ladder in the United States since around the year 2000.

4.2.2 Academic peer effects

Academic peer effects are spillover effects in the production of knowledge. They arise if learning in higher education depends on the quality of one's peers in higher education. As human capital raises earnings, such peer quality ultimately affects the monetary payoffs from higher education as well. Peer effects are conceptually very similar to externalities. Individuals at the participation margin can gain from participation, but have a negative effect on all existing participants as participation reduces average peer quality. Hence, if individuals do not take this effect into account, then peer effects lead to too much participation.[21]

We interpret type as quality. We suppose that the benefit of higher education depends on your own quality θ, as before, but also on the average quality of your peers, being $E[\theta|\theta \geq \theta^\circ]$ with θ° the equilibrium cutoff. With slight abuse of notation, we write the benefit of higher education now as $b(\theta, E[\theta|\theta \geq \theta^\circ])$, with b differentiable and strictly increasing. Let $\tilde{b}(\theta^\circ) = b(\theta^\circ, E[\theta|\theta \geq \theta^\circ])$ and note that $\tilde{b}$ strictly increases with θ°. If individuals correctly anticipate the equilibrium cutoff, then an individual with type θ will study if

$$a(\theta) + b(\theta, E[\theta|\theta \geq \theta^\circ]) - c(\theta) \geq a(\theta),$$

leading to a unique equilibrium cutoff type θ° satisfying

$$\tilde{b}(\theta^\circ) - c(\theta^\circ) = 0. \tag{8}$$

To find the optimal set of participants, the average payoff in society can be written as

$$\int a(\theta)f(\theta)d\theta + \int_{\geq\theta^\star} (b(\theta, E[\theta|\theta \geq \theta^\star]) - c(\theta))f(\theta)d\theta.$$

The first-order condition is

$$(\tilde{b}(\theta^\star) - c(\theta^\star))f(\theta^\star) - p(\theta^\star) = 0,$$

[21] We assume that students cannot sort according to ability, say, there is only one higher education institution such that higher quality students cannot leave for another institution if too much lower quality peers enter. Although no mobility is clearly unrealistic, the main conclusion does not change if one relaxes this assumption. Increases in participation will then lead to resorting at all levels and thereby lower average peer quality throughout the distribution of participants in higher education.

with

$$p(\theta^\star) = \int_{\geq\theta^\star} b_2'(\theta, E[\theta|\theta \geq \theta^\star]) \frac{dE[\theta|\theta \geq \theta^\star]}{d\theta^\star} f(\theta) d\theta.$$

The first term $(\tilde{b}(\theta^\star) - c(\theta^\star))f(\theta^\star)$ is the net benefit of participation at the margin, while the second term $-p(\theta^\star)$ captures the total negative effect of the increase in participation at the margin on all current participants because average peer quality decreases. The first-order condition implies that the first term $\tilde{b}(\theta^\star) - c(\theta^\star)$ is strictly positive. Combined with equation (8), we get

$$\tilde{b}(\theta^\star) - c(\theta^\star) > \tilde{b}(\theta^\circ) - c(\theta^\circ).$$

Because $\tilde{b} - c$ is strictly increasing, we obtain $\theta^\star > \theta^\circ$. So, academic peer effects indeed imply inefficiently high participation. The situation is essentially the opposite of the externality case; there is a negative impact on other individuals that is not taken into account in the participation decision.

Empirical research on academic peer effects in higher education is limited.[22] The quality of college roommates is found to have a positive effect on academic performance (Sacerdote, 2001; Winston and Zimmerman, 2004; Sacerdote, 2011), but these effects are very modest in magnitude. The intensity of the interaction plays a role, however. A unique experiment at an air force academy shows that peer effects are larger if peers eat, study, and work together in small and isolated groups (Carrell et al., 2009). A follow-up study underlines the complex dynamics behind peer effects in higher education. Carrell et al. (2013) use the findings of the initial study to experimentally construct peer groups to maximize the outcomes of low-ability peers at the air force academy. However, the program produced opposite results: the formation of homogeneous subgroups within each peer group led to decreases in the performance of low-ability peers.

There also exists recent evidence from randomized experiments in the Netherlands on peer effects at the university level. Booij et al. (2015) find that low-ability students benefit from sorting, while high-ability students are unaffected. The result is driven by a negative impact of the standard deviation in quality on low-ability students. Feld and Zölitz (2014) find a very low positive effect of the average GPA, which is relatively strongest for low-ability students. Grades are higher when there are less students at both extremes of the ability distribution. Hence, both studies indicate that there can be

[22]There is no direct evidence of the role of academic peer effects in the decision to participate in higher education, hence we mainly focus on peer effects in relation to achievement.

a negative impact from having more students of very high ability. A recent study from Luppino and Sander (2015) suggests that the more competitive nature of higher education might explain such results. They find that attending a campus with higher quality peers in science studies reduces the chance of obtaining a science degree, lowers GPA, and decreases graduation rates. On the other hand, while Feld and Zölitz (2014) finds that the addition of students from the low end of the ability distribution is harmful, the findings from Booij et al. (2015) indicate that the addition of another low-ability student can have benefits for other low-ability students, which would go against the theoretical assumption that we made before that marginal increases in participation provide a 'negative externality' through peer quality. However, more evidence on such effects at the higher education level is needed.

4.2.3 Peer pressure

Peer interactions may also result from peer pressure; choices of peers that directly influence utility. For example, conformity within groups of high school students can play a role in the decision to participate in higher education. If conformity is a behavioural failure—i.e., if individuals regret later in life that they did conform as adolescents—then intervention can be justified.

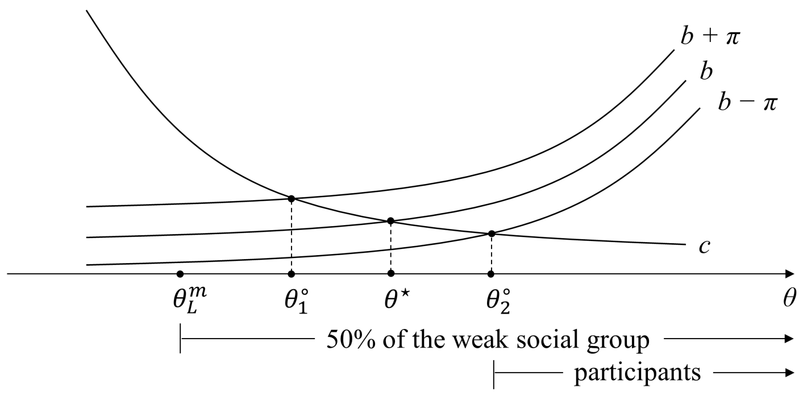
Conformity in higher education arises if the willingness to participate increases with the proportion of individuals of your social group that participates (and vice-versa for non-participation). We model conformity by adding an extra payoff, say $\pi > 0$, if the choice of an individual turns out to be a majority choice within his or her social group.[23] The other payoff functions (a, b, and c) remain the same and, in particular, they do not differ among social groups. However, the distribution of ability types can be different among the different groups. Social groups will be indexed by i and are exogenous for ease of exposition.

Let $\mathbb{1}[\cdot]$ be a dummy indicator that equals 1 if the statement between brackets is true (and zero otherwise) and let P_i=1-$F(\theta)_i$ be the proportion of participants in social group i. An individual of social group i with type θ will choose to study if

$$a(\theta) + b(\theta) + \pi.\mathbb{1}[P_i^\circ \geq 0.5] - c(\theta) \geq a(\theta) + \pi.\mathbb{1}[P_i^\circ < 0.5],$$

[23] Conformity and identity are related to each other. Identity can be modelled as the loss in utility if one deviates from the social group ideal; see Akerlof and Kranton (2000). If the group ideal is endogenously defined by the majority choice in the group, then conformity and identity coincide.

Figure 4: Peer pressure



Notes: The figure denotes the benefit (b) and costs (c) of higher education, as well as alternative benefit functions where we either assume a positive or a negative additional payoff (π) depending on whether participation conforms with the majority decision of the peer group or not. These payoffs are depicted across student type θ. θ_L^m represents the median ability level of a social group with relatively low ability.

or equivalently, if

$$b(\theta) - c(\theta) + \pi\{\mathbb{1}[P_i^\circ \geq 0.5] - \mathbb{1}[P_i^\circ < 0.5]\} \geq 0.$$

Figure 4 plots four curves, the original benefit (*b*) and cost *(c)* functions, as well as two alternative benefit functions $b \pm \pi$.

The efficient cutoff for each social group is the same and satisfies $b(\theta^\star) - c(\theta^\star) = 0$. The two cutoffs θ_1° and θ_2° satisfy $b(\theta_1^\circ) - c(\theta_1^\circ) - \pi = 0$ and $b(\theta_2^\circ) - c(\theta_2^\circ) + \pi = 0$ and are potential equilibrium cutoffs.

Consider an arbitrary social group i and call θ_i^m the median ability, defined by $F_i(\theta_i^m) = 0.5$. In Figure 4, we consider a social group with a low median ability, being $\theta_L^m < \theta_1^\circ$. Think of high school students at the end of secondary education sorted in a weak track. We show that θ_1° cannot be an equilibrium outcome. If individuals in the weak group anticipate a majority that chooses to study, then $\mathbb{1}[P_i^\circ \geq 0.5] - \mathbb{1}[P_i^\circ < 0.5] = 1$ and indeed θ_1° would result. But the anticipation turns out to be false, because the resulting participation rate is equal to $1 - F(\theta_1^\circ)$, which is smaller than 0.5 (because $\theta_L^m < \theta_1^\circ$). In contrast, the other potential equilibrium cutoff θ_2° is self-fulfilling.

We conclude that conformity in weak social groups leads to too little participation (compare θ_2° with $\theta^\star$). Following the same reasoning, it is easy to show that conformity in social groups with a high median ability ($\theta_2^\circ < \theta_H^m$) leads to too much participation because now θ_1° is the self-fulfilling equilibrium cutoff that lies to the left of $\theta^\star$. For groups with intermediate ability ($\theta_1^\circ < \theta_I^m < \theta_2^\circ$), both under- and overparticipation can occur, leading to an inefficiency in both cases.

There is empirical evidence that conformity in secondary education can hinder academic preparation for college. Gaviria and Raphael (2001) identify peer-group influences for school dropout. An increase in the mean dropout among peers by 1% is associated with a 0.16% increase in individual dropout probability and this effect is robust across a wide range of specifications. ‘Acting white,’ a phenomenon that refers to the fact that good study ethics are socially undesirable in some minority youth cultures, serves as another potential mechanism. Fryer and Torelli (2010) present a rare empirical study on this topic. They find that in-group status monotonically increases with academic performance for white students. For black and especially Hispanic students, this positive relationship reverses above a certain threshold. It implies that too much diligence at school may lead to expulsion from the social group. The fear of expulsion and identity loss may lead to underperformance in high school, and, as a consequence, students could be insufficiently prepared to start in higher education.

Conformity also occurs in higher education. Research shows that peer pressure is important for social outcomes such as drinking, smoking, substance use, and crime among students.[24] There is little evidence however of the influence of peer pressure on decisions in higher education. Stinebrickner and Stinebrickner (2008) show that college roommates that bring video games have a negative impact on grades, because they lower study effort. Italian evidence shows that one additional classroom peer who chooses economics over business increases the chance of others also choosing economics as a major by 1 percentage point (De Giorgi et al., 2009). As a consequence, students might not always choose the major that best reflects their academic ability or individual preferences.

[24]For example, Sacerdote (2011) indicates that having a roommate in college who smokes increases the chance of smoking by around 5 to 10 percentage points. Estimates are even larger for binge drinking.

5 First-best failures

5.1 From first-best to second-best

The failures discussed in previous sections lead to inefficiencies. Except for uninsurable risks, these inefficiencies are reflected in differences between the equilibrium and the optimal participation level in higher education. Table 2 provides an overview of the different failures. For each failure we mention how the equilibrium and optimal cutoff types θ° and $\theta^\star$ are defined and indicate the predicted consequences for the corresponding equilibrium and optimal participation levels $P^\circ = 1 - F(\theta^\circ)$ and $P^\star = 1 - F(\theta^\star)$.

Table 2: Overview of the different failures

	$d^\circ(\theta) =$	$d^\star(\theta) =$	Part. gap
No failures	$b(\theta) - c(\theta)$	$b(\theta) - c(\theta)$	$P^\circ = P^\star$
Market failures			
- uninsurable risk	$b(\theta) - c(\theta) - \frac{\rho}{2}\Delta\sigma$	$b(\theta) - c(\theta)$	$P^\circ < P^\star$
- credit constraints	$min\{b(\theta) - c(\theta), m(\theta) - c(\theta)\}$	$b(\theta) - c(\theta)$	$P^\circ < P^\star$
- externalities	$b(\theta) - c(\theta)$	$b(\theta) + e(\theta) - c(\theta)$	$P^\circ < P^\star$
- signalling	$\Delta w(\theta) - c(\theta)$	$b(\theta) - c(\theta)$	$P^\circ > P^\star$
Behavioural failures			
- misprediction	$\hat{b}(\theta) - \hat{c}(\theta)$	$b(\theta) - c(\theta)$	$P^\circ < P^\star$
- positional goods	$b(\theta) + s - c(\theta)$	$b(\theta) - c(\theta)$	$P^\circ > P^\star$
- academic peer effects	$\tilde{b}(\theta) - c(\theta)$	$\tilde{b}(\theta) - c(\theta) - p(\theta)$	$P^\circ > P^\star$
- peer pressure	$b(\theta) - c(\theta) \pm \pi$	$b(\theta) - c(\theta)$	both

Notes: The table shows the equilibrium (d°) and the optimality ($d^\star$) functions that determine the equilibrium and optimal cutoff types θ° and $\theta^\star$, as well as the expected participation gap (actual participation versus optimal participation). These results are given for the benchmark case with no failures as well as the different cases where either market or behavioral failures occur.

In a first-best world, policy makers are able to use lump-sum transfers. A lump-sum transfer is a tax (or subsidy) that is non-distortive because the tax amount cannot be changed by changing behaviour. If lump-sum transfers were available, then the second fundamental welfare theorem tells us that efficiency and equity issues can be separated. A simple solution to the market failures presented in Table 2 is to provide a participation subsidy (or tax) equal to $d^\star(\theta^\star) - d^\circ(\theta^\star)$.[25] It will shift the equilibrium cutoff θ° towards the optimal cutoff $\theta^\star$. Non-distortive lump-sum transfers can be used to

[25] In case of uncertainty, it is better to offer insurance directly.

raise the necesary funds for the subsidy and to deal with distributional issues if deemed necessary.

The availability of lump-sum transfers presupposes that the individual types θ are observable. This is an unrealistic assumption. In case of education and earnings, a type reflects intelligence and diligence which are both private information. Therefore, second-best policies, e.g., subsidizing education financed by taxes on earnings, have to be used in reality. This raises new issues.

5.2 Education subsidies and redistribution as "Siamese twins"

Education and redistribution policies are connected in several ways. In the words of Bovenberg and Jacobs (2005), "Redistribution and education subsidies are Siamese twins." As a consequence, education and redistribution policies should not be designed in isolation. We postpone the optimal joint design of education and redistribution policies to section 6. In this section, we prepare the discussion and focus on three new issues: the distortive effect of subsidies and taxes, the possible existence of (so-called) fiscal externalities, and the regressivity of education subsidies.

5.2.1 The distortive effect of education subsidies and redistribution on participation

Taxes on labour income can distort the participation decision (Trostel, 1993), but subsidizing higher education encourages participation and may therefore counteract the previous distortion (Trostel, 1996; Bovenberg and Jacobs, 2005). To see this, consider a simple linear tax and subsidy scheme in a one-period model without uncertainty; we will discuss some extensions—tax-deductible costs, uncertainty, non-linear taxation, and capital income taxation—later on. Labour income taxation consists of a lump-sum tax T and a proportional tax rate τ applied to (ex-post) gross income $y(\theta)$, being either $a(\theta)$ or $a(\theta) + b(\theta)$. The education subsidy scheme consists of a lump-sum S and a proportional subsidy rate ς applied to the cost of education $c(\theta)$.[26]

An individual with type θ will participate in higher education if

$$-T + (1-\tau)(a(\theta) + b(\theta)) + S - (1-\varsigma)c(\theta) \geq -T + (1-\tau)a(\theta),$$

[26]Benefits and costs of higher education are assumed to be monetary here and thus taxable. Note that the conclusions here do not change if one splits up benefits and costs into a monetary and non-monetary part.

or equivalently, if

$$b(\theta) + (\varsigma c(\theta) + S - \tau b(\theta)) \geq c(\theta).$$

Let θ° be the type that is indifferent between participating or not; thus, the cutoff type θ° implicitly satisfies

$$b(\theta^\circ) + (\varsigma c(\theta^\circ) + S - \tau b(\theta^\circ)) = c(\theta^\circ). \tag{9}$$

The change in participation $P^\circ = 1 - F(\theta^\circ)$ for a small change in the different policy instruments $X = S, T, \varsigma, \tau$ is given by $\frac{\partial P^\circ}{\partial X} = -f(\theta^\circ)\frac{\partial \theta^\circ}{\partial X}$, with

$$\frac{\partial \theta^\circ}{\partial S} = -\frac{1}{(1-\tau)b'(\theta^\circ) - (1-\varsigma)c'(\theta^\circ)} < 0, \tag{10}$$

$$\frac{\partial \theta^\circ}{\partial T} = 0, \tag{11}$$

$$\frac{\partial \theta^\circ}{\partial \varsigma} = -\frac{c(\theta^\circ)}{(1-\tau)b'(\theta^\circ) - (1-\varsigma)c'(\theta^\circ)} < 0, \tag{12}$$

and

$$\frac{\partial \theta^\circ}{\partial \tau} = \frac{b(\theta^\circ)}{(1-\tau)b'(\theta^\circ) - (1-\varsigma)c'(\theta^\circ)} > 0, \tag{13}$$

(assuming $\varsigma, \tau < 1$). Hence, participation decreases with the tax rate, but increases with the different subsidy instruments; the lump-sum income tax has no effect on participation. In other words, whereas the tax on labour income discourages participation, the subsidy on education encourages it and counteracts therefore the effect of taxation.

In addition, the different distortions may interact (Bovenberg and Jacobs, 2005). For example, the educational subsidy S may be less distortive, the higher the tax rate τ. The cross-derivative of participation with respect to the subsidy S and the tax rate τ is

$$\frac{\partial^2 P^\circ}{\partial S \partial \tau} = -f'(\theta^\circ)\frac{\partial \theta^\circ}{\partial S}\frac{\partial \theta^\circ}{\partial \tau} - f(\theta^\circ)\frac{\partial^2 \theta^\circ}{\partial S \partial \tau}. \tag{14}$$

If the sign is positive, then education subsidies are more effective at higher levels of taxation. Although a positive sign can be derived under reasonable assumptions, we do not provide further details here.[27] The main message is simply that subsidies and tax rates may interact in complicated ways.

There exist several extensions in the literature that are worth discussing. First, Boskin (1975) and Heckman (1976) argue that taxes do not necessarily discourage investment in human capital, if costs are made tax-deductible. To see this, suppose that educational costs are not subsidized, but made completely tax-deductible, then an individual with type θ will participate if

$$-T + (1-\tau)(a(\theta) + b(\theta) - c(\theta)) \geq -T + (1-\tau)a(\theta),$$

which reduces indeed to equation (1) again. Note that this argument holds only if taxes are linear and the educational costs in c are purely monetary.

Second, if there is uncertainty that cannot be insured, then the income tax rate will provide insurance against income risk (Levhari and Weiss, 1974; Eaton and Rosen, 1980; Krebs, 2003). With uncertainty, equation (13) changes to

$$\frac{\partial \theta^{\circ}}{\partial \tau} = \frac{b(\theta^{\circ}) - \rho(1-\tau)(\sigma_{\beta}^{2} - \sigma_{\alpha}^{2})}{(1-\tau)b'(\theta^{\circ}) - (1-\varsigma)c'(\theta^{\circ})},$$

with ρ the coefficient of (constant) absolute risk aversion and $\sigma_{\beta}^{2} - \sigma_{\alpha}^{2}$ the difference in income risk between participants and non-participants.[28] So, if individuals are risk averse and if participants face a higher income risk compared to non-participants, then the negative impact of taxation on participation will indeed be lower and can even become positive.

Third, allowing for progressivity in tax rates may lead to different outcomes. Trostel (1993) already emphasizes that the negative impacts of taxation will be more substantial when we assume progressive taxation. Heckman et al. (2008) find that moving from a progressive to a flax tax rate would increase enrolment by 19 percentage points in the United States. Holter (2014) finds that introducing the progressivity of the Danish tax system in the United States would decrease enrolment by 8 percentage points. Although substantial, both simulations do not include the general equilibrium effect of enrolment

[27] If the benefit of higher education increases concavely with type ($b'' < 0$) and if the cost of higher education decreases convexly ($c'' > 0$), then the cross-derivative $\frac{\partial^2 \theta^{\circ}}{\partial S \partial \tau}$ is always negative (proof is in appendix). So if $f'(\theta^{\circ}) \geq 0$ (e.g., in case of a uniform distribution or if more than half of the population participates for a normal distribution) or if $f'(\theta^{\circ})$ is not too small, then the cross-derivative $\frac{\partial^2 P^{\circ}}{\partial S \partial \tau}$ will be positive.

[28] We derive this result in equation (24) of the appendix, with $\eta = 0$.

on wages. Heckman et al. (1998) show that the net impact on enrolment becomes negligible once general equilibrium effects are included.

Fourth, capital income taxation discourages investment in financial assets, but makes investment in human capital relatively more interesting. Jacobs and Bovenberg (2010) consider a two-period life cycle model to allow for the cumulation of different assets, human and financial, over time. Because labour income taxation encourages individuals to replace human by financial assets, capital income taxation counteracts this replacement and can therefore be a useful supplement to labour taxation as well.

5.2.2 Fiscal 'externalities'

Taxes on labour income may be distortive, but they can also lower the budgetary cost of subsidizing higher education. Subsidies are clearly a cost for society, but, as they encourage participation in higher education, they also lead to higher earnings and thus to higher tax revenues. Because increasing subsidies to higher education never makes someone worse off directly, they can even turn out to be Pareto improving if these extra subsidies do not reduce tax revenues (Findeisen and Sachs, 2014a). Some authors call the tax revenue effect caused by subsidizing higher education a fiscal externality. Although there is no real externality here, we will stick to this term later on.[29]

To derive the effect of subsidies on tax revenues, consider again the linear tax and subsidy scheme.[30] The per-capita tax revenue is equal to

$$R = T + \tau \bar{a} + \int_{\geq \theta^\circ} (\tau b(\theta) - S - \varsigma c(\theta)) f(\theta) d\theta,$$

with $\bar{a} = \int a(\theta) f(\theta) d\theta$ and θ° the cutoff type as defined before. A marginal change in the lump-sum subsidy S or in the subsidy rate ς causes a marginal change in tax revenues equal to

$$\frac{\partial R}{\partial S} = \tau b(\theta^\circ) \frac{\partial P^\circ}{\partial S} - [(S + \varsigma c(\theta^\circ)) \frac{\partial P^\circ}{\partial S} + P^\circ], \tag{15}$$

and

$$\frac{\partial R}{\partial \varsigma} = \tau b(\theta^\circ) \frac{\partial P^\circ}{\partial \varsigma} - [(S + \varsigma c(\theta^\circ)) \frac{\partial P^\circ}{\partial \varsigma} + \bar{c}^\circ P^\circ], \tag{16}$$

[29] An externality in the classical sense could still occur, if there are different government layers such that the subsidies of one layer affect the revenues of another layer.

[30] For ease of exposition, we neglect the distortive effect of taxes and subsidies on labour and education effort. We come back to it in Section 6.

with $\frac{\partial P^\circ}{\partial X}$ for $X = S, \varsigma$ as defined before and $\bar{c}^\circ = \int_{\geq \theta^\circ} c(\theta) f(\theta) d\theta / P^\circ$ the average educational cost of participants in higher education.

In words, given a positive tax rate on labour income, increasing subsidies on higher education will be Pareto improving if the marginal increase in tax revenues through increased participation (the first term in each equation) exceeds the marginal cost of the subsidy (the second term between squared brackets). The marginal cost of the subsidy consists of two parts: providing the full subsidy to the new participants (the first part within the squared brackets) and providing the marginal subsidy to the existing participants (the second part).[31] These formulas have to be evaluated at the margin. Hence, because lower ability types have both lower benefits (b) and higher costs (c), a Pareto improvement is less likely to happen at higher levels of subsidization.

Some recent papers have tried to quantify the magnitude of fiscal externalities. Findeisen and Sachs (2014a) estimate that a one dollar increase in subsidies to higher education in the United States increases future discounted tax revenues with at least 0.7 dollars on average and up to 1.65 dollars if targeted to children of low income parents. Holter (2014) finds that applying the Danish subsidy scheme to the United States would lead to an increase in net tax revenues in a partial equilibrium setting. Lawson (2015) shows that the fiscal externalities of increasing financial aid in the United States are large enough to generate a Pareto improvement. Although robust to many specifications, this result is very sensitive to general equilibrium and spillover effects.

5.2.3 Education subsidies can be regressive

The use of general taxes on labour income to subsidize higher education can be regressive. The reasoning is that those who do not participate in higher education have a lower average lifetime income and thereby co-finance the education of 'richer' participants. Interestingly, this concern has been raised as an objection against subsidizing education from general taxation both by those left (Marx, 1875) and right (Friedman, 1955) of the political spectrum.[32] To see this, we must introduce a more general welfare approach encompassing efficiency and equity.

Let $w(\theta)$ denote the welfare weight attached to an individual with type θ and assume without loss of generality that these weights add up to one, i.e., $\bar{w} = \int w(\theta) f(\theta) d\theta = 1$.

[31] The cost of the subsidy is overestimated from a welfare point of view, because the subsidy also benefits students and is not wasted. This will become more clear in Section 6.

[32] Marx focuses on basic, rather than higher education.

Efficiency requires the weight function w to be positive everywhere ($w > 0$), while equity imposes higher weights for individuals with lower types as they will have a lower income ($w' < 0$). Let $u(\theta)$ be the utility of an individual with type θ and denote average utility by $\bar{u} = \int u(\theta)f(\theta)d\theta$. Welfare is equal to the sum of weighted utilities, being

$$\int w(\theta)u(\theta)f(\theta)d\theta = \bar{u} + \int (w(\theta) - \bar{w})(u(\theta) - \bar{u})f(\theta)d\theta.$$

The decomposition shows that welfare can be written as the sum of an efficiency term (the average utility) and an equity term (the covariance between weights and utilities). In particular, equity goes up if those with a higher welfare weight would obtain a higher utility.

The simplest way to see that subsidizing education can be inequitable is to increase the lump-sum subsidy S and the lump-sum tax T simultaneously in a budget neutral way. Let

$$\bar{w}^\circ = \int_{\geq\theta^\circ} w(\theta)f(\theta)d\theta/P^\circ \tag{17}$$

be the average weight of participants in higher education. We show in the appendix that subsidizing higher education decreases equity if $1 - \bar{w}^\circ > 0$; i.e. if the welfare benefit of one euro spent exceeds the cost of 1 euro. This condition is always true given the assumptions $\bar{w} = 1$ and $w' < 0$.

Although obvious in a simple setting where everybody pays a tax T to finance a subsidy S for the better off only, it becomes less obvious in general. For example, one could think of increasing subsidies for higher education in a selective way, e.g., for those around the participation margin, and financing the subsidy by progressively increasing the taxes for the rich only. Our overview of the empirical results reflect this general complexity.

There is an on-going debate on the possible regressivity of educational subsidies. The first empirical studies on the topic take a cross-sectional point of view, meaning that they address the distributional impact of subsidies and taxes on the income of parents of prospective students. Hansen and Weisbrod (1969) show that educational subsidies are regressive in the Californian higher education system. Pechman (1970) contests the thesis and initiates the so-called Hansen-Weisbrod-Pechman debate. In his overview of the empirical literature, Barbaro (2006) concludes that regressivity is far from clear. More recent studies consider the distributional impact of educational subsidies in a longitudinal view, i.e., looking at lifecycle earnings of graduates and non-graduates. Gruske

(1994) and Johnson (1984) find a regressive impact in Germany and a neutral or mildly progressive impact in the United States, respectively.

Besides the complexity of the research question itself, general equilibrium and non-linear spillover effects may also play a role in the ambiguous findings of the empirical literature. General equilibrium effects occur if subsidies to higher education lead to a relatively higher supply of skilled workers and a lower college premium in equilibrium. Because they imply more equal wages between skilled and unskilled workers, general equilibrium effects may counteract the alleged regressivity of education subsidies (Johnson, 1984). Similarly, if graduates have positive wage spillover effects that affect the wages of non-graduates more than the wages of other graduates, then again more equal wages between skilled and unskilled workers may result, reducing the regressive nature of subsidies. We discuss both effects in more detail.

There is clear evidence that participation in higher education produces general equilibrium effects. Several authors provide evidence that the significant rise in the college premium in the 1980s can be largely attributed to a sharp deceleration in the relative supply of college workers; see, e.g., Autor et al. (2008). Heckman et al. (1998) show that the effects of tuition subsidies on enrolment as well as earnings inequality are almost completely offset by general equilibrium effects. On the other hand, Lee (2005) finds that such effects are present but relatively minor; the identified impact of a 1% increase in tuition on enrolment are 1.27% and 1.05% in the partial and general equilibrium case, respectively.[33] Acemoglu (1998) illustrates that long-run general equilibrium effects can be very different from short-run general equilibrium effects. Increases in the supply of skilled labour induces directed skill-biased technological change, which can offset short-term decreases in the college wage premium.

There is also some empirical evidence for non-linear spill-over effects. Moretti (2004) reports evidence that a one percentage point increase in the labour force share of college graduates in a metropolitan area raises wages of high-school dropouts by 1.9%, those of high school graduates by 1.6%, and those of college graduates by 0.4%. Ciccone and Peri (2006) provide a contrasting view, arguing that standard methods can mistakenly identify general equilibrium effects resulting from the change in the relative supply of high-skilled workers for externalities. Correcting for the labour force skill composition leads to estimates that are lower and not statistically significant. However,

[33]Lee (2005) provides some explanations for the difference between his findings and those of Heckman et al. (1998). Most prominently, the model in Lee (2005) allows for those that take up extra education as a result of the subsidy to remain in the low-skilled sector.

the estimates from Ciccone and Peri (2006) are relatively imprecise. In fact, the level of precision would not have identified the estimates of Moretti (2004) as statistically significant either.

6 Optimal education and redistribution policies

Section 5 reminds us that, without lump-sum transfers, there are no easy solutions to remedy the market and behavioural failures discussed in Sections 3 and 4. In particular, education and redistribution policies turn out to be closely linked and should therefore not be dealt with separately. In this final section we focus on the optimal joint design of education and redistribution policies. First, we analyze the classical case, i.e., the optimal design of subsidies to higher education financed by general earnings taxation. Afterwards, we discuss the potential role of innovative financing modes like income-contingent loans and graduate taxes.

6.1 Education subsidies paid by earnings taxes

For reasons of tractability, we cannot model all failures that we introduced before. For simplicity, we assume that each ability type θ has a net effect on society equal to $e(\theta)$. Recall Table 2. The net effect is best interpreted as the sum of the different relevant market and behavioural failures that drive a positive or negative wedge between what is individually optimal and socially desirable.[34] addition, we add uncertainty and incomplete insurance to the model. These are essential features of the participation decision in higher education that cannot be captured by the overall net effect.[35]

To explain the basic intuition, we consider first linear tax and subsidy schemes; non-linear schemes will be discussed later on. As before, the participation decision is explicitly modelled and the distortions of labour taxes and education subsidies on participation are therefore automatically included. To capture also the distortions on labour and study effort, we follow Alesina and Giuliano (2009) and assume that distortions are quadratic in the tax and subsidy rate, i.e., $\frac{1}{2}d_\tau\tau^2$ and $\frac{1}{2}d_\varsigma\varsigma^2$, with d_τ and d_ς measuring the magnitude of the distortion.[36] Both terms are expressed in per-capita revenue, so, they will directly enter the budget constraint of the social planner as a budgetary cost.

[34]We assume the net effect to be increasing ($e' > 0$), but not necessarily positive.

[35]We assume a constant coefficient of absolute risk aversion $\rho > 0$.

[36]See, e.g., Saez et al. (2012) for a recent review of the literature on the (distortive) impact of earnings taxation on taxable income.

We use the same welfare approach as before, but, because uncertainty is present, we replace the utility of each type by its expected utility. So, efficiency ($w > 0$) will now also guarantee that insurance will be approved of, while equity ($w' < 0$) allows for solidarity from individuals with better life-time prospects to those with worse prospects.

Three first-order conditions are obtained (for S, τ, and ς), while the budget constraint fixes the poll tax (T). The appendix provides all derivations.[37] We discuss the interpretation of each first-order condition in detail later on.

The different instruments serve different purposes. The lump-sum educational subsidy S is used only to guarantee an optimal participation level in higher education from a societal point of view, i.e., taking the net effect of participants on society (e) and equity aspects (via w) into account. The tax rate τ and the subsidy rate ς redistribute either earnings or study costs from individuals with better life-time prospects to individuals with worse life-time prospects. In addition, the tax rate and subsidy rate will also partially insure workers against income risk (captured by α and β as before) or participants against study cost risk (captured by γ), respectively. Although the role of the tax and subsidy rates τ and ς is more extensive compared to the lump-sum subsidy S, note also that they come at a higher cost because they do not only distort the participation decision, but also labour and study effort.

6.1.1 The lump-sum educational subsidy S

The welfare effect of the lump-sum subsidy S is equal to

$$\frac{\partial W}{\partial S} = \frac{\partial R}{\partial S} + e(\theta^{\circ})\frac{\partial P^{\circ}}{\partial S} + \bar{w}^{\circ} P^{\circ}, \tag{18}$$

with $\frac{\partial R}{\partial S}$ the fiscal revenue effect described by equation (15). So, the welfare effect is equal to the pure fiscal revenue effect augmented by two additional terms. The first additional term is the (positive or negative) net effect of the new participants on societal welfare. To understand the second additional term, note that the welfare cost of one euro subsidy to all existing participants is less than one euro: it is equal to one minus the average welfare weight of the current participants. Because only the pure cost aspect is captured by the fiscal revenue effect $\frac{\partial R}{\partial S}$ the second additional term has to be interpreted as a welfare correction because subsidies to existing participants are not pure waste.

[37] In the appendix we also add a graduate tax rate η to be discussed later on in section 6.2. So, the results mentioned here in section 6.1 are derived by setting $\eta = 0$.

If there are no net effects on society ($e = 0$) and if the planner is Rawlsian ($\bar{w}^{\circ} = 0$), then the additional terms are zero and welfare maximization coincides with revenue maximization. In all other cases, the welfare effect is likely to be larger than the pure revenue effect (if one assumes $e > 0$).

The welfare effect in equation (18) can be used in at least three ways. First, in terms of marginal reform, subsidies should be set higher if the welfare effect is positive at the margin and vice-versa. Second, in terms of optimal design, the optimum requires the marginal welfare effect to be zero, i.e., at the optimum $\theta^{\star}$ we must have

$$e(\theta^{\star}) + b(\theta^{\star}) - c(\theta^{\star}) = \frac{1 - \bar{w}^{\star}}{\frac{\partial P^{\star}}{\partial S}\frac{1}{P^{\star}}}, \tag{19}$$

using equations (9) and (15). The left-hand side of equation (19) is the net social benefit of participation at the margin, while the right-hand side is the ratio of the welfare effect of a subsidy to existing participants and the semi-elasticity (the percentage change in participation caused by a small change in subsidy). The optimal participation level mixes therefore equity and efficiency considerations. If only efficiency matters ($w \to 1$ everywhere) or if the semi-elasticity of participation is perfectly elastic ($\frac{\partial P^{\circ}}{\partial S}\frac{1}{P^{\circ}} \to +\infty$), then the right-hand side becomes zero. The subsidy S should then be set to guarantee the first-best optimal level of participation (such that $e(\theta^{\star}) + b(\theta^{\star}) - c(\theta^{\star}) = 0$). If equity matters and if the elasticity is less than perfectly elastic, then the right-hand side is positive and there will be less participation in second-best compared to first-best.

Third, if one assumes a welfare-maximizing planner, then equation (19) must hold and each component of the formula can be derived from knowledge of the other components. The inverse optimum exercise, first introduced in Ahmad and Stern (1984) for indirect taxation, derives the implicit welfare weight structure, but it could also be interesting to retrieve, for example, the implicit net effect at the margin.

6.1.2 The earnings tax rate τ

If the lump-sum subsidy S is optimally set, then the welfare effect of the tax rate τ is equal to

$$\left.\frac{\partial W}{\partial \tau}\right|_{S^{\star}} = net\ redistributive\ payoff + insurance\ payoff - cost\ of\ distortion.$$

The net redistributive payoff is equal to

$$net\, redistributive\, payoff \; = -cov(w,y) - (1-\bar{w}^{\circ})P^{\circ}b(\theta^{\circ}) > 0. \qquad (20)$$

The first part is the pure redistributive payoff of income taxation. It is equal to the marginal increase in equity which is, in our model, measured by the covariance between weights and gross incomes, being

$$cov(w,y) = \int (w(\theta)-\bar{w})(y(\theta)-\bar{y})f(\theta)d\theta < 0,$$

with $y(\theta)$ the ex-post gross income (either $a(\theta)$ or $a(\theta)+b(\theta)$). The second part is the welfare cost of distorting participation in higher education.[38] More precisely, it is defined as the welfare cost (captured by $1-\bar{w}^{\circ}$) to restore the participation distortion caused by the tax rate by exactly countersubsidizing all participants in higher education (a proportion P°) with a subsidy $S = b(\theta^{\circ})$. Although not immediately clear from the terms in equation (20), the net redistributive payoff is positive (as shown in the appendix).

The insurance payoff is the value of the decrease in income risk caused by taxation, i.e.,

$$insurance\, payoff = \rho(1-\tau)\sigma_y^2 > 0,$$

with ρ the inequality aversion parameter, $1-\tau$ the change in income risk caused by taxation, and $\sigma_y^2 = (1-P^{\circ})\sigma_\alpha^2 + P^{\circ}\sigma_\beta^2$ the ex-post gross income risk.

Finally, both payoffs are counteracted by the marginal cost of taxation which is simply equal to

$$cost\ of\ distortion = d_\tau \tau > 0.$$

In the absence of equity motives ($w \to 1$), there is only an insurance payoff and a marginal welfare cost of distorting labour. The tax rate will then be in between no insurance ($\tau \to 0$) and full insurance ($\tau \to 1$) depending on whether the cost of distortion is very high ($d_\tau \to \infty$) or negligeable ($d_\tau \to 0$).

[38] We subtract the welfare cost of the distortion here, rather than adding it to the cost of the distortion later on. Although the latter may seem more logical at first sight, our choice will turn out to be useful later on to sign unambiguously the (net) redistributive payoff of the subsidy rate in section 6.1.3.

6.1.3 The educational subsidy rate ς

The welfare effect of the subsidy rate ς, conditional on an optimal lump-sum subsidy $S^{\star}$, consists of the same three parts, so

$$\left.\frac{\partial W}{\partial \varsigma}\right|_{S^{\star}} = net\ redistributive\ payoff + insurance\ payoff - cost\ of\ distortion,$$

but the terms are different.

The net redistributive payoff of the subsidy rate is

$$net\ redistributive\ payoff = cov(w, z) + (1 - \bar{w}^{\circ})P^{\circ}c(\theta^{\circ}) > 0.$$

The first part is again the equity increase of the subsidy rate, now measured by the covariance between weights and gross study costs, i.e.,

$$cov(w, z) = \int (w(\theta) - \bar{w})(z(\theta) - \bar{z})f(\theta)d\theta,$$

with $z(\theta)$ the ex-post study cost (either 0 or $c(\theta)$). The second part is the welfare gain of distorting participation. To see this, note that slightly increasing the subsidy rate encourages participation, so we can now lower the subsidy S with an amount $c(\theta^{\circ})$ for all participants (a proportion P° of the population) to restore participation. Thus, there is now a welfare gain equal to $(1 - \bar{w}^{\circ})P^{\circ}c(\theta^{\circ})$ to be added. Again we show in the appendix that the net redistributive payoff is positive.

The insurance payoff depends on the inequality aversion ρ, the change in study risk $1 - \varsigma$, and the study risk itself $P^{\circ}\sigma_{\gamma}^{2}$ and is equal to

$$insurance\ payoff = \rho(1 - \varsigma)P^{\circ}\sigma_{\gamma}^{2} > 0.$$

Finally, the cost of the distorting study effort is

$$cost\ of\ distortion = d_{\varsigma}\varsigma > 0.$$

In the absence of equity motives ($w \to 1$), the insurance motive again leads to a degree of insurance in between no insurance ($\varsigma \to 0$) and full insurance ($\varsigma \to 1$) depending on whether the cost of the distortion d_{ς} is high ($d_{\varsigma} \to \infty$) or low ($d_{\varsigma} \to 0$).

6.2 Income-contingent loans and graduate taxes

Up to now we have discussed loans and general taxation as two possible ways to finance the cost of higher education. But there exist alternative financing schemes such as income-contingent loans and graduate taxes. We compare the different financing modes along several dimensions and look at the consequences of these differences for some of the problems that we discussed before. We also characterize an optimal graduate tax and conclude with a discussion of real-world applications.

6.2.1 Financing modes and financing features

Loans (LOA) and general taxation (GET) are different in several aspects: (1) a tax is paid by everyone in society, while a loan is paid by the students only, (2) a tax is compulsory, while a loan is contracted on a voluntary basis, (3) a tax payment depends on your future income, while a loan installment is fixed, and (4) the total tax is not capped from above, while a loan is, i.e. you never pay back more than the cost of higher education (including interest).[39]

Alternative financing schemes have been advocated in the literature, most prominently, income-contingent loan (ICLs) and graduate taxes (GRTs).[40] The nature of each scheme is a combination of the properties of both the classical loan and the general tax. First, ICLs and GRTs have in common with a general tax that the repayment depends on future income: if income is low in some particular period, the required periodic repayment is temporarily lower and vice-versa for when income is high. Second, ICLs and GRTs have in common with a loan that mainly students, not society, pay back the costs of higher education. ICLs and GRTs also exhibit some differences. GRTs (like general taxes) are compulsory and not capped, while ICLs (like loans) are not compulsory and capped.[41]

Table 3 summarizes the different financing modes (in rows) and their properties (in columns). The last column indicates how the default risk—the risk that a student pays back less than the incurred cost of higher education over the lifetime—is treated. It can

[39] The legal status of a tax and a loan is different as well, which can be relevant when, for example, students move abroad after graduation. We do not discuss this aspect here.

[40] See e.g. Barr (2001); Palacios (2004); Chapman (2006) for extensive overviews. The term 'graduate tax' is slightly misleading, because it does not require graduation. 'Participation tax' or 'study tax' would be more accurate, as it is usually conceived as a tax on study credits.

[41] This is how the literature generally characterizes these schemes, although one could theoretically also imagine a GRT that is made voluntary and/or capped as well.

be shared by everyone in society, or it can be pooled, e.g., by student cohort.[42] In case of a risk-sharing arrangement, the cost of default is borne by everyone in society and we write therefore '(mainly) students' in the first column, as students do not exclusively pay the costs of higher education in this case.

Table 3: Overview of the different financing modes

	who pays	compulsory	income-dependent	capped	default risk
GET	all taxpayers	yes	yes	no	shared
GRT	(mainly) students	yes	yes	no	shared/pooled
LOA	(mainly) students	no	no	yes	shared/pooled
ICL	(mainly) students	no	yes	yes	shared/pooled

Notes: GET = general tax, GRT = graduate tax, LOA = loan, and ICL = income-contingent loan

To illustrate a stylized version of ICLs and GRTs, we assume for simplicity that the cost of higher education is the same for all students and that the income in each repayment period is constant over the lifetime (but not necessarily the same for different students). Figure 5 shows the *required* lifetime repayment (vertical axis) as a function of lifetime income (horizontal axis) for a linear graduate tax scheme and an income-contingent loan.[43] The cost of higher education for a certain type of individual is indicated on the vertical axis. The corresponding dotted horizontal line can be interpreted as the required lifetime repayment under a classical loan.[44]

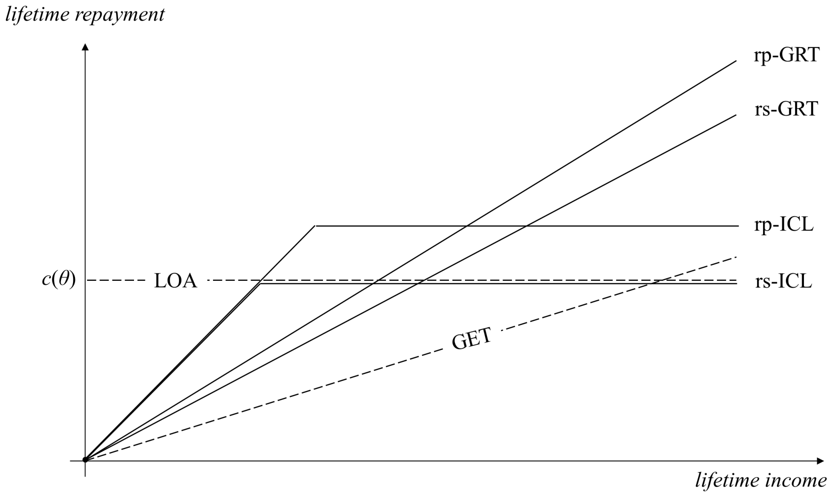
Because ICL's are capped, the maximum cost is either equal to the cost of higher education (in case of a risk-sharing income contingent loan, denoted rs-ICL) or somewhat higher (in case of a risk-pooling income contingent loan, denoted rp-ICL) to cover the cost of default. GRTs are not capped, but according to the same logic the tax rate of a risk-pooling GRT must be higher than the tax rate of a risk-sharing GRT to cover the default costs. Because the total revenues raised by the GRT and ICL scheme (for the same default risk arrangement) in Figure 5 are kept the same, the tax rate of a GRT scheme will be initially lower than the repayment rate of an ICL, but higher afterwards

[42] A classical loan can be risk-sharing, e.g., if provided or guaranteed by the government.

[43] Although irrelevant to compare the different schemes in theory, notice that most ICLs in practice do not require repayment if your income falls below a certain threshold.

[44] We implicitly assume that either the loan repayment is guaranteed by the government or the loan is provided by the government and is of the risk-sharing type.

Figure 5: Repayments under different financing schemes



Notes: The figure shows the required lifetime repayment of the costs of higher education as a function of lifetime income, separately for a general tax (GET), classical loan (LOA), income-contingent loan (ICL) and graduate tax (GRT). For the ICL and the GRT, we distinguish within each scheme between risk-sharing (rs) and risk-pooling (rp) arrangements.

when the repayment rate of an ICL drops to zero. As a matter of comparison, we also indicate the linear tax rate of a general tax (GET) scheme by a dotted line. Because everyone in society contributes, the general tax rate will be lower than the graduate tax rates.[45]

6.2.2 Consequences of the different financing features

The features of the different financing modes have consequences for some of the problems that we discussed before. We look at the features of ICLs and GRTs and link them to insurance and consumption smoothing, moral hazard, adverse selection, participation and redistribution, and fairness.

Income dependence allows for insurance (ex-ante) and consumption smoothing (ex-post); see Chapman (2006). First, income dependence is a common feature of ICLs and GRTs for low income levels. Because default is mainly a problem of low income, this common feature offers insurance against default risk from an ex-ante point of view.[46]

[45]Note that this refers here to the share of the general tax income that is used to finance higher education.

[46]Default is expected to be lower for GRTs compared to ICLs, because the tax rate of a GRT is initially lower than the repayment rates of ICLs.

Second, if we compare a capped ICL scheme with an uncapped GRT scheme (for the same default risk arrangement) in Figure 5, then it should be clear that low income students pay back more in the ICL scheme while high income students pay back less. As such, the distribution of disposable incomes—incomes minus repayments—of students will be more compressed under a GRT compared to an ICL; see Garcia-Penalosa and Wälde (2000). A GRT scheme offers therefore more insurance against income risk from an ex-ante point of view. Third, income dependence leads to consumption smoothing from an ex-post point of view, because disposable incomes will be more balanced over the lifetime of a student. Similarly, GRTs lead to a more compressed distribution of disposable incomes for individuals over time, and thus more consumption smoothing compared to ICLs.

An immediate drawback of income dependence is a higher potential for moral hazard (especially compared to loans) because it could induce students to choose jobs with lower labour hours or lower effort after their studies; see, e.g., Nerlove (1975). For example, Feldman (1976) conducts a range of simulations and finds that allowing for moral hazard in the context of an ICL reduces the labour supply (measured in weeks worked) of medical specialists by 6.6%. Even if moral hazard is a feature of all income-dependent schemes, ICLs and GRTs may be less vulnerable compared to GETs. Findeisen and Sachs (2014b) show that if labour supply elasticities also depend on education levels, then ICLs and GRTs can raise the same tax revenues with less distortions compared to GETs, because the former allow for differentiation of tax rates for individuals with different levels of education.

ICLs are not compulsory and can therefore be vulnerable to adverse selection if they pool the default risk. The Yale Plan, a risk-pooling ICL that was introduced at Yale University in 1972, offers a clear example; see Nerlove (1975). Adverse selection led to an unexpectedly high initial default rate, which in turn led to substantial increases in the repayments of students that remained in the scheme. This induced some of those students to avoid repaying as well, which further increased repayments. Eventually, the scheme was abandoned. Although the impact of adverse selection can be self-destructive, it is not always the case. Del Rey and Racionero (2014) show that, under specific conditions, a risk-pooling ICL can in theory be sustainable without coercion.

The impact of financing schemes on participation, especially of disadvantaged groups in society, is a politically sensitive issue in many societies. We discuss therefore participation here, although we like to stress that more participation should not necessarily be interpreted as better from a welfare point of view.[47] We have seen in section 3.1 that the participation decision under uncertainty depends on the expected net benefits of higher education, the variance of net benefits as a measure of risk, and the level of risk aversion. The expected costs under a GRT scheme will be rather similar to the expected costs under an ICL, ceteris paribus.[48] A GRT scheme offers more insurance than an ICL scheme, because the variance of disposable incomes is likely to be smaller under a GRT. Taken together, GRTs would lead to more participation compared to ICLs. The comparison between risk-sharing and risk-pooling schemes is more complicated because risk-sharing leads to lower expected costs for students, but also higher variances in disposable income, ceteris paribus. If students are sufficiently risk averse, then participation will be higher under risk-pooling and vice-versa (Del Rey and Racionero, 2010). Risk aversion also plays a role when comparing GRTs and ICLs to general taxation (GET). The tax rate is lower under a GET scheme, which implies lower expected costs, but also less insurance. Again, if students are sufficiently risk averse, then participation can be higher under GRTs and ICLs compared to GETs and vice-versa (Del Rey and Racionero, 2010).

Lower expected costs for students may lead to higher participation in higher education, but also to a less progressive distribution of income over the life cycle. Indeed, the lower the expected costs for students, the more the non-students have to contribute to the financing of higher education. Because non-students are likely to be poorer on average over their lifetime, lower costs for students result in a less progressive redistribution of income over the lifetime. For example, risk-sharing schemes have lower expected costs for students and can thus be expected to be more perverse compared to risk-pooling; see, e.g., Garcia-Penalosa and Wälde (2000).

With respect to fairness, ICLs and a GRTs provide more equal opportunities (especially compared to classical loans) by alleviating credit constraints. If we compare ICLs and GRTs, some claim that a GRT is less fair than an ICL, because the total (ex-post) repayment may (greatly) exceed the incurred cost of higher education and can potentially also be far above the obtained return (Chapman, 2006). This is labeled by Barr (2001) as the 'Mick Jagger problem.'[49]

[47] See Del Rey and Racionero (2010) for a comparison of the participation level under the different financing schemes relative to the efficient participation level.

[48] If the impact of moral hazard is similar and if everyone participates in non-compulsory schemes (ICLs), then the budget constraint implies the same expected cost under both ICLs and GRTs.

[49] It refers to the lead singer of the Rolling Stones who briefly studied at LSE and who would have therefore faced massive repayments under a graduate tax.

6.2.3 An optimal graduate tax

Up to now we have discussed the features of alternative financing modes and their implications. Because adverse selection is a potential problem of ICLs, we focus here on graduate taxes.[50] The crucial question is the following: suppose that society can already tax earnings (via T and τ) and subsidize higher education (via S and ς), can a graduate tax further enhance welfare?

Before discussing the results, we want to clarify some details of the welfare problem. First, graduate tax revenues will be used to cover the subsidies to higher education, rather than the full costs. Covering full costs of higher education is a special case of subsidization (choosing $S = 0$ and $\varsigma = 1$), so, the specification we use is more flexible from a welfare point of view. Second, we start with a simple linear graduate tax scheme here and we will discuss more general non-linear tax schemes later on. Because there is already a lump-sum subsidy S to higher education, it does not make sense to also introduce a lump-sum graduate tax. A linear graduate tax scheme consists therefore only of a graduate tax rate, denoted by η. Third, we must clarify whether the graduate tax scheme will be risk-sharing or risk-pooling. If the graduate tax is risk-pooling, then two separate budget constraints have to be met: a graduate tax constraint requiring that the graduate tax revenues (raised via η and net of its distortion costs like default) cover the subsidies to higher education (spent via S and ς and including the distortion costs of ς) and an earnings tax constraint requiring that the earnings tax revenues (raised via T and τ and net of the distortion costs of τ) are non-negative. If the graduate tax is risk-sharing, then the same two budget constraints have to be met, but the distortion cost of the graduate tax shifts from the graduate tax constraint to the earnings tax constraint. Because two separate budget constraints are always less flexible and thus more constraining compared to the corresponding joint budget constraint—all tax revenues (raised via T, τ, and η) must cover all subsidies (spent via S and ς)—we obtain immediately that risk-pooling and risk-sharing schemes are welfare inferior compared to adding the joint budget constraint to the welfare maximization problem.

We analyze a planner that maximizes welfare using earnings taxes (T and τ), subsidies (S and ς), and a graduate tax rate (η) as instruments and subject to the joint budget constraint. The role of S, T, τ, and ς is the same as before. The role of the graduate tax

[50] One could think of compulsory ICLs. Besides unnatural to think of a compulsory loan from a legal point of view, it should also be clear from Figure 5 that a non-linear graduate tax can mimick a compulsory ICL.

rate is similar to the role of the earnings tax rate. Given an optimal lump-sum subsidy to higher education, the welfare effect of the graduate tax rate can also be written as

$$\left.\frac{\partial W}{\partial \eta}\right|_{S^\star} = net\ redistributive\ payoff + insurance\ payoff - cost\ of\ distortion,$$

but in contrast to the optimal earnings tax rate (see Section 6.1.2), the (net) redistributive payoff and the insurance payoff are among graduates only. To put it differently, the graduate tax rate regulates additional redistribution and insurance among graduates.

Although linear graduate tax schemes can further enhance welfare, this is not true anymore if we allow for non-linear tax schemes in our model. Because the marginal cost of distortion is zero at the laisser-faire, we must have $\tau^\star, \eta^\star > 0$ at the optimum. Recall also that there will be a cutoff type $\theta^\star$ that separates graduates from non-graduates at the optimum. As a consequence, the tax rate for all individuals with a lower type (and thus lower earnings) will be $\tau^\star$, while those with a higher type face a tax rate $\tau^\star + \eta^\star$. The combination of a linear earnings tax and a graduate tax scheme can therefore be exactly mimicked by a progressive piece-wise linear earnings tax scheme (and no graduate tax scheme).

If we allow tax schemes to be non-linear, then graduate tax schemes are superfluous in our model. Yet, we must stress that the following three assumptions are crucial: (1) there are no framing effects, in particular, the salience of income taxes and graduate taxes is the same, (2) there exists an income level that separates graduates and non-graduates, (3) the graduate tax rate cannot depend on the study cost. While there is little evidence to support or falsify the first assumption, the second and third assumption do not hold for sure. Hence, a graduate tax can better target the repayment of study costs to participants and avoids therefore the perverse redistributive effects of general taxation.

6.2.4 Real-world applications

There are no real-world examples of graduates taxes, but there exist several past and present examples of income-contingent loans. Most of the existing schemes today are risk-sharing ICLs; see Chapman (2006) for an overview.[51] The empirical literature on the implications of these schemes is still thin as most schemes are implemented rela-

[51] Besides the terminated Yale Plan, there currently exists a risk-pooling ICL scheme in Hungary; see Berlinger (2009) for details.

tively recently. The existing evidence suggests that ICLs do not have adverse impacts on the participation levels of disadvantaged groups, relative to schemes with lower student contributions. Andrews (1999) finds that the proportion of low-income students in Australian higher education did not change after strong reductions in the generosity of the financing scheme in 1997. Chapman and Ryan (2002) find that participation rates did not decrease for students from any family wealth subgroup, and neither the introduction of the ICL nor the strong changes in 1997 led to changes in the socio-economic mix of participants in higher education. Dearden et al. (2011) conduct a similar analysis for the United Kingdom. They find that the introduction of the ICL scheme in 1998, which coincided with strong (income-dependent) increases in tuition, only lowered the participation rate of students with high parental income. The 2006 reforms, which led to further increased and uniform tuition fees, did not impact the participation rate of any of the subgroups, as the impact of higher tuition was completely offset by the expansion of grants and ICLs.

The previous studies for Australia and the United Kingdom are evaluations of (new) ICL policies versus (old) general tax schemes. Because the participation effects are small, even negligible in most studies, prospective students can be expected to be risk averse with respect to higher education enrolment (otherwise participation should decline as the expected costs increase when replacing GETs by ICLs).[52] Alternatively, the low participation impact might also be a consequence of a psychological effect, e.g., students fail to rationally incorporate the increase in expected costs. Regardless of the specific results of these studies, one should be aware that evaluating ICLs vs. GETs on the basis of participation rates does not allow for any conclusions with respect to the welfare effects of these schemes.

7 Conclusion

This paper has reviewed the economics of financing higher education with a focus on why intervention could be desirable and how it can be done. We highlight the main messages to conclude our review.

We have specified different arguments why intervention in higher education could be desirable. Section 3 discusses the traditional market failures (uninsurable risk, credit

[52]This is provided that credit constraints are not too high in the original case, which appears to apply in these cases, as the old system had no tuition fees combined with maintenance grants for students from poor families.

constraints, externalities, and signalling). Some of these failures, externalities and signalling in particular, seem plausible a priori, but turn out to be hard to prove in a convincing way. Although we would not go as far as to discard such arguments completely, caution (and more empirical research) is needed in our view when it comes to emphasizing these failures as justification for intervention. Other market failures, like credit constraints, have become more important over time in countries with relatively high private contributions for higher education attendance, such as the U.S. and the U.K. The increasing importance of credit constraints could spread to other countries as well, especially to those that have recently opted to shift a larger share of the costs of higher education towards students and their parents without accompanying measures to alleviate credit constraints. Section 4 discusses several behavioural failures that could arise via misprediction and social interactions (positional goods, academic peer effects, and peer pressure). Because behavioural economics is a relatively recent field, the empirical knowledge is still thin, and this is especially true in the context of higher education. Still, the available evidence indicates that these behavioral failures play a role in decision-making processes for higher education and therefore need to be considered when designing policies in higher education.

Sections 5 and 6 analyze how interventions in higher education can be executed. Section 5 discusses the problems that may arise when ideal lump-sum transfers are not feasible for redistribution. The use of second-best instruments, like a classical income tax to finance subsidies for higher education, may indeed lead to several complications (distortions, perverse redistribution, and fiscal externalities). Although each of these complications is often taken for granted, the empirical evidence is not overwhelming. Section 6 looks at the joint design of education and redistribution policies. Central in the design of a classical tax-subsidy scheme is the trade-off between on the one hand redistribution and insurance benefits and on the other hand the distortion costs. Section 6 also discusses more innovative schemes like income-contingent loans and graduate taxes. Compared to general taxation, such alternative financing modes are often praised for better targeting the costs to students, while keeping the consumption-smoothing and insurance properties of general taxation. Recent evidence suggests that income-contingent loans do not constrain students from low income families to participate in higher education. Although no real world examples of graduate taxes exist, the same is likely to be true for that particular financing scheme as well.

Our review not only updates the academic literature, but also comes at a time of prevalent reforms across educational systems worldwide. The reforms in higher education finance appear to have been largely inspired by pragmatic considerations: how to save on government expenditures in higher education, usually by shifting more of the costs to students and their parents, while at the same time trying to limit the impact on participation in higher education. Although revenues and participation are no doubt important dimensions of higher education, neither a higher participation level nor a lower budget deficit are necessarily better from a welfare point of view. Additionally, rather than reducing the budget deficit, the resulting savings could also be reinvested in structural and long-term educational policies. Early childhood investment, for example, has the potential to improve the life prospects of children from disadvantaged families and to provide true democratic access to higher education in the long-run.

References

Acemoglu, D. (1998). Why do new technologies complement skills? Directed technical change and wage inequality. *Quarterly journal of economics 113*(4), 1055–1089.

Ahmad, E. and N. Stern (1984). The theory of reform and indian indirect taxes. *Journal of Public Economics 25*(3), 259–298.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *Quarterly journal of Economics 115*(3), 715–753.

Alesina, A. F. and P. Giuliano (2009). Preferences for redistribution. National Bureau of Economic Research, Discussion Paper no. 14825.

Andrews, L. (1999). The effect of HECS on access. Research Report, Department of Education, Employment, Training and Youth Affairs, Canberra.

Angrist, J., D. Lang, and P. Oreopoulos (2009). Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics 1*(1), 136–163.

Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies 29*(3), 155–173.

Arrow, K. J. (1973, July). Higher education as a filter. *Journal of Public Economics 2*(3), 193–216.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008). Trends in U.S. wage inequality: Revising the revisionists. *Review of Economics and Statistics 90*(2), 300–323.

Avery, C. (2010). The effects of college counseling on high-achieving, low-income students. National Bureau of Economic Research Working paper no. 16359.

Avery, C. and C. M. Hoxby (2004). Do and should financial aid packages affect students' college choices? In C. Hoxby (Ed.), *College choices: The economics of where to go, when to go, and how to pay for it*, pp. 239–302. Chicago: University of Chicago Press.

Barbaro, S. (2006). *Equity and efficiency considerations of public higher education*. New York: Springer Science & Business Media.

Barr, N. (2001). *The Welfare State as Piggy Bank: Information, Risk, Uncertainty, and the Role of the State*. Oxford, UK: OUP Oxford.

Beaudry, P., D. A. Green, and B. M. Sand (2013). The great reversal in the demand for skill and cognitive tasks. National Bureau of Economic Research Working paper no. 16359.

Bedard, K. (2001). Human capital versus signaling models: university access and high school dropouts. *Journal of Political Economy 109*(4), 749–775.

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

Belman, D. and J. S. Heywood (1991). Sheepskin effects in the returns to education: An examination of women and minorities. *The Review of Economics and Statistics 73*(4), 720–724.

Belzil, C. and J. Hansen (2004). *Research in Labor Economics: Accounting for worker well-being*, Chapter Earnings Dispersion, Risk Aversion and Education, pp. 335–358. Emerald group publishing, Bingley, United Kingdom.

Belzil, C. and M. Leonardi (2007). Can risk aversion explain schooling attainments? Evidence from Italy. *Labour Economics 14*(6), 957–970.

Berlinger, E. (2009). An efficient student loan system: Case study of Hungary. *Higher Education in Europe 34*(2), 257–267.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012). The role of application assistance and information in college decisions: Results from the h&r block fafsa experiment. *The Quarterly Journal of Economics 127*(3), 1205–1242.

Bhuller, M., M. Mogstad, and K. G. Salvanes (2011). Life-cycle bias and the returns to schooling in current and lifetime earnings. NHH Dept. of Economics Discussion Paper no. 4.

Booij, A. S., E. Leuven, and H. Oosterbeek (2015). Ability peer effects in university: Evidence from a randomized experiment. IZA Discussion Paper, no. 8769.

Boskin, M. J. (1975). Notes on the tax treatment of human capital. National Bureau of Economic Research, Working Paper no. 116.

Bovenberg, A. L. and B. Jacobs (2005). Redistribution and education subsidies are Siamese twins. *Journal of Public Economics 89*(11), 2005–2035.

Caetano, G. S., M. Palacios, and H. A. Patrinos (2011). Measuring aversion to debt: an experiment among student loan candidates. World Bank Policy Research Working Paper Series, no. 5737.

Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue, and M. Rabin (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". *University of Pennysylvania Law Review 151*(??), 1211–1254.

Cameron, S. V. and J. J. Heckman (2001, June). The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy 109*(3), 455–99.

Carneiro, P., K. Hansen, and J. J. Heckman (2003, May). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review 44*(2), 361–422.

Carneiro, P. and J. J. Heckman (2002, October). The evidence on credit constraints in post-secondary schooling. *Economic Journal 112*(482), 705–734.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011). Estimating marginal returns to education. *American Economic Review 101*(6), 2754–2781.

Carrell, S. E., R. L. Fullerton, and J. E. West (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics 27*(3), 439–464.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica 81*(3), 855–882.

Chan, D. and D. F. Cochrane (2008). Paving the way: How financial aid awareness affects college access and success. Institute for College Access and Success.

Chapman, B. (2006). Income contingent loans for higher education: International reforms. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 2, pp. 1435–1503. Amsterdam: Elsevier.

Chapman, B. and C. Ryan (2002). Income-contingent financing of student charges for higher education: Assessing the australian innovation. *The University of Wales Journal of Education 11*(1), 64–81.

Chen, S. H. et al. (2001). Is investing in college education risky? State University of New York at Albany, Department of Education Discussion Papers no. 01-09.

Chevalier, A., C. Harmon, I. Walker, and Y. Zhu (2004). Does education raise productivity, or just reflect it? *The Economic Journal 114*(499), F499–F517.

Chowdry, H., C. Crawford, L. Dearden, A. Goodman, and A. Vignoles (2013). Widening participation in higher education: Analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 176*(2), 431–457.

Ciccone, A. and G. Peri (2006). Identifying human-capital externalities: Theory with applications. *The Review of Economic Studies 73*(2), 381–412.

Clark, A. E., P. Frijters, and M. A. Shields (2008). Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles. *Journal of Economic literature 46*(1), 95–144.

Connolly, S. and A. Munro (1999). *Economics of the public sector*. Prentice Hall.

Conti, G., J. J. Heckman, and S. Urzua (2010, May). The education-health gradient. *American Economic Review: Papers and Proceedings 100*(2), 1–5.

Cunha, F., J. J. Heckman, L. J. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Chapter 12, pp. 697–812. Amsterdam: North-Holland.

Currie, J. and E. Moretti (2003, November). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics 118*(4), 1495–1532.

Cutler, D. M. and A. Lleras-Muney (2008). Education and health: Evaluating theories and evidence. In R. F. Schoeni, J. S. House, G. A. Kaplan, and H. Pollack (Eds.), *Making Americans Healthier: Social and Economic Policy as Health Policy*, pp. 29–60. New York: Russell Sage Foundation.

De Giorgi, G., M. Pellizzari, and S. Redaelli (2009). Be as careful of the company you keep as of the books you read: Peer effects in education and on the labor market. National Bureau of Economic Research, Working Paper no. 14948.

Dearden, L., E. Fitzsimons, and G. Wyness (2011). The impact of tuition fees and support on university participation in the UK. IFS Working Papers W11/17.

Dee, T. S. (2004). Are there civic returns to education? *Journal of Public Economics 88*(9), 1697–1720.

Del Rey, E. and M. Racionero (2010). Financing schemes for higher education. *European Journal of Political Economy 26*(1), 104–113.

Del Rey, E. and M. D. M. Racionero (2014). Choosing the type of income-contingent loan: risk-sharing versus risk-pooling. IEB Working Paper.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature 47*(2), 315–72.

Duncan, G. J. and J. Brooks-Gunn (1997). Income effects across the life span: Integration and interpretation. In G. Duncan and J. Brooks-Gunn (Eds.), *Consequences of Growing Up Poor*, pp. 596–610. New York: Russell Sage Foundation.

Dynarski, S. M. and J. E. Scott-Clayton (2006). The cost of complexity in federal student aid: Lessons from optimal tax theory and behavioral economics. *National Tax Journal 59*(2), 319–356.

Eaton, J. and H. S. Rosen (1980). Taxation, human capital, and uncertainty. *The American Economic Review 70*(4), 705–715.

Falch, T. and H. Oosterbeek (2011). Financing lifelong learning: Funding mechanisms in education and training. EENEE Analytical Report no. 10.

Feld, J. and U. Zölitz (2014). On the nature of peer effects in academic achievement. University of Gothenburg: Working papers in Economics No. 596.

Feldman, R. (1976). Some more problems with income-contingent loans: The case of medical education. *The Journal of Political Economy 84*(6), 1305–1311.

Findeisen, S. and D. Sachs (2014a). Designing efficient college and tax policies. Unpublihsed Manuscipt.

Findeisen, S. and D. Sachs (2014b). Education and optimal dynamic taxation: The role of income-contingent student loans. Unpublished Manusript. Available at SSRN: http://papers.ssrn.com/abstract=1957170.

Fricke, H. (2013). Tuition fees and student achievement-evidence from differential tuition fees. In *IV Workshop on Economics of Education-Higher Education Financing Policy: Schemes and Effects, Barcelona*.

Friedman, M. (1955). *The role of government in education*. East Lansing: Rutgers University Press.

Fryer, R. G. and P. Torelli (2010). An empirical analysis of 'acting white'. *Journal of Public Economics 94*(5), 380–396.

Garcia-Penalosa, C. and K. Wälde (2000). Efficiency and equity effects of subsidies to higher education. *Oxford Economic Papers 52*(4), 702–722.

Garibaldi, P., F. Giavazzi, A. Ichino, and E. Rettore (2012). College cost and time to complete a degree: Evidence from tuition discontinuities. *Review of Economics and Statistics 94*(3), 699–711.

Gautier, P. A., G. J. Van den Berg, J. C. Van Ours, and G. Ridder (2002). Worker turnover at the firm level and crowding out of lower educated workers. *European Economic Review 46*(3), 523–538.

Gaviria, A. and S. Raphael (2001). School-based peer effects and juvenile behavior. *Review of Economics and Statistics 83*(2), 257–268.

Gibson, J. (2000). Sheepskin effects and the returns to education in New Zealand: do they differ by ethnic groups? *New Zealand Economic Papers 34*(2), 201–220.

Griffin, P. and A. Cox Edwards (1993). Rates of return to education in Brazil: Do labor market conditions matter? *Economics of Education Review 12*(3), 245–256.

Grimes, P. W. (2002). The overconfident principles of economics student: An examination of a metacognitive skill. *The Journal of Economic Education 33*(1), 15–30.

Grodsky, E. and M. T. Jones (2007). Real and imagined barriers to college entry: Perceptions of cost. *Social Science Research 36*(2), 745–766.

Groot, W. and H. Oosterbeek (1992). Optimal investment in human capital under uncertainty. *Economics of Education Review 11*(1), 41–49.

Gross, J. P., O. Cekic, D. Hossler, and N. Hillman (2009). What matters in student loan default: A review of the research literature. *Journal of Student Financial Aid 39*(1), 19–29.

Grossman, M. (2000). The human capital model. In A. J. Culyer and J. P. Newhouse (Eds.), *Handbook of Health Economics*, Volume 1, pp. 347–408. Amsterdam: Elsevier.

Grossman, M. (2006). Education and nonmarket outcomes. In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 1, Chapter 10, pp. 577–633. Amsterdam: Elsevier.

Gruske, K.-D. (1994). Verteilungseffekte deroffentlichen hochschulfinanzierung in der bundesrepublik deutschland-personale inzidenz im querschnitt und langsschnitt. *Bildung, Bildungsfinanzierung und Einkommensverteilung 2*, 71–147.

Hansen, W. L. and B. A. Weisbrod (1969). The distribution of costs and direct benefits of public higher education: the case of California. *Journal of human resources 4*(2), 176–191.

Harmon, C., H. Oosterbeek, and I. Walker (2003). The returns to education: Microeconomics. *Journal of Economic Surveys 17*(2), 115–141.

Harrison, G. W., M. I. Lau, and M. B. Williams (2002). Estimating individual discount rates in Denmark: A field experiment. *American economic review 92*(5), 1606–1617.

Heckman, J. J. (1976, August). A life-cycle model of earnings, learning, and consumption. *Journal of Political Economy 84*(4, Part 2), S11–S44. Journal Special Issue: Essays in Labor Economics in Honor of H. Gregg Lewis.

Heckman, J. J., J. E. Humphries, S. Urzúa, and G. Veramendi (2011). The effects of educational choices on labor market, health, and social outcomes. Unpublished manuscript, University of Chicago, Department of Economics.

Heckman, J. J., L. J. Lochner, and C. Taber (1998, May). General-equilibrium treatment effects: A study of tuition policy. *American Economic Review 88*(2), 381–386.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2008, Spring). Earnings functions and rates of return. *Journal of Human Capital 2*(1), 1–31.

Heller, D. E. (1997). Student price response in higher education: An update to Leslie and Brinkman. *Journal of Higher education 68*(6), 624–659.

Hilger, N. G. (2014). Parental credit constraints and college attendance: Evidence from timing of fathers layoffs. Unpublished Manuscript.

Holmlund, H., M. Lindahl, and E. Plug (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature 49*(3), 615–651.

Holter, H. A. (2014). Accounting for cross-country differences in intergenerational earnings persistence: the impact of taxation and public education expenditure. Available at SSRN: http://papers.ssrn.com/abstract=1978261.

Horn, L. J., X. Chen, and C. Chapman (2003). Getting ready to pay for college. US Department of Education/Institute of Education Sciences: NCES.

Hungerford, T. and G. Solon (1987, February). Sheepskin effects in the returns to education. *Review of Economics and Statistics 69*(1), 175–177.

Jabbar, H. (2011). The behavioral economics of education: New directions for research. *Educational Researcher 40*(9), 446–453.

Jacobs, B. and A. L. Bovenberg (2010). Human capital and optimal positive taxation of capital income. *International Tax and Public Finance 17*(5), 451–478.

Jaeger, D. A. and M. E. Page (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *Review of Economics and Statistics 78*(4), 733–740.

Johnson, C. and C. Montmarquette (2011). Loan aversion among Canadian high school students. CIRANO-Scientific Publications 2011s-67.

Johnson, G. E. (1984). Subsidies for higher education. *Journal of Labor Economics 2*(3), 303–318.

Johnston, G. (2004). Healthy, wealthy and wise? A review of the wider benefits of education. New Zealand Treasury Working Paper 04/04.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics 17*(2), 303–316.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–291.

Keane, M. P. and K. I. Wolpin (2001, November). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review 42*(4), 1051–1103.

Kleven, H. J. and W. Kopczuk (2008). Transfer program complexity and the take up of social benefits. National Bureau of Economic Research Working paper no. 14301.

Krebs, T. (2003). Human capital risk and economic growth. *The Quarterly Journal of Economics 118*(2), 709–744.

Krueger, A. B. and M. Lindahl (2001). Education for growth: Why and for whom? *Journal of Economic Literature 39*, 1101–1136.

Lang, K. and D. Kropp (1986). Human capital versus sorting: the effects of compulsory attendance laws. *The Quarterly Journal of Economics 101*(3), 609–624.

Lange, F. and R. Topel (2006). The social value of education and human capital. *Handbook of the Economics of Education 1*, 459–509.

Lawson, N. (2015). Liquidity constraints, fiscal externalities and optimal tuition subsidies. Unpublished Manuscipt. Available at SSRN: http://papers.ssrn.com/abstract=2585062.

Layard, R. and G. Psacharopoulos (1974). The screening hypothesis and the returns to education. *The Journal of Political Economy 82*(5), 985–998.

Lee, D. (2005). An estimable dynamic general equilbrium model of work, schooling, and occupational choice. *International Economic Review 46*(1), 1–34.

Leuven, E., H. Oosterbeek, and B. van der Klaauw (2010). The effect of financial rewards on students' achievement: Evidence from a randomized experiment. *Journal of the European Economic Association 8*(6), 1243–1265.

Levhari, D. and Y. Weiss (1974). The effect of risk on the investment in human capital. *The American Economic Review 64*(6), 950–963.

Lindahl, M., M. Palme, S. Sandgren Massih, and A. Sjögren (2012). The intergenerational persistence of human capital: an empirical analysis of four generations. IZA Discussion Paper, no. 6463.

Lochner, L. J. and A. Monge-Naranjo (2011). The nature of credit constraints and human capital. *The American Economic Review 101*(6), 24–87.

Lochner, L. J. and E. Moretti (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review 94*(1), 155–189.

Luppino, M. and R. Sander (2015). College major peer effects and attrition from the sciences. *IZA Journal of Labor Economics 4*(1), 1–23.

Machin, S., O. Marie, and S. Vujić (2011). The crime reducing effect of education. *The Economic Journal 121*(552), 463–484.

Marx, K. (1875). Critique of the Gotha programme. *Karl Marx and Frederick Eneels, Collected Works 24*.

Matus-Grossman, L. and S. Gooden (2002). Opening doors: Students' perspectives on juggling work, family, and college. Manpower Demonstration Research Group, New York.

Mazzeo, C. (2007). Private lending and student borrowing: a primer. In F. M. Hess (Ed.), *Footing the tuition bill: The new student loan sector*, pp. 74–97.

McMahon, W. W. (2004). The social and external benefits of education. In G. Johnes and J. Johnes (Eds.), *International handbook on the economics of education*, pp. 211–259. Cheltenham, UK: Edward Elgar.

Meghir, C. and L. Pistaferri (2004, January). Income variance dynamics and heterogeneity. *Econometrica 72*(1), 1–32.

Milligan, K., E. Moretti, and P. Oreopoulos (2004, August). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics 88*(9-10), 1667–1695.

Moretti, E. (2004). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *American Economic Review 94*(3), 656–690.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1991). The allocation of talent: Implications for growth. *The Quarterly Journal of Economics 106*(2), 503–30.

Nelson, R. R. and E. S. Phelps (1966, March). Investment in humans, technological diffusion, and economic growth. *American Economic Review 56*(1/2), 69–75.

Nerlove, M. (1975). Some problems in the use of income-contingent loans for the finance of higher education. *The Journal of Political Economy 83*(1), 157–183.

Nowell, C. and R. M. Alston (2007). I thought I got an A! Overconfidence across the economics curriculum. *The Journal of Economic Education 38*(2), 131–142.

Oosterbeek, H. (1992). Study duration and earnings: A test in relation to the human capital versus screening debate. *Economics letters 40*(2), 223–228.

Palacios, M. (2004). *Investing in human capital: a capital markets approach to student funding*. New York: Cambridge University Press.

Pallais, A. (2015). Small differences that matter: Mistakes in applying to college. *Journal of Labor Economics 33*(2), 493–520.

Patrinos, H. A. (1996). Non-linearities in the returns to education: sheepskin effects or threshold levels of human capital? *Applied Economics Letters 3*(3), 171–173.

Pechman, J. A. (1970). The distributional effects of public higher education in California: A review article. *Journal of Human Resources 5*(3), 361–370.

Peters, E., D. Västfjäll, P. Slovic, C. Mertz, K. Mazzocco, and S. Dickert (2006). Numeracy and decision making. *Psychological science 17*(5), 407–413.

Pollmann-Schult, M. (2005). Crowding-out of unskilled workers in the business cycle: Evidence from West Germany. *European Sociological Review 21*(5), 467–480.

Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature 36*(1), 11–46.

Rabin, M. (2002). A perspective on psychology and economics. *European economic review 46*(4), 657–685.

Romer, P. (1993). Idea gaps and object gaps in economic development. *Journal of Monetary Economics 32*(3), 543–573.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics 116*(2), 681–704.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In E. Hanushek, S. Machin, and L. Woessman (Eds.), *Handbook of the Economics of Education*, Volume 3 of *Handbooks in Economics*, Chapter 4, pp. 249–277. Amsterdam: North-Holland.

Saez, E., J. Slemrod, and S. H. Giertz (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature 50*(1), 3–50.

Sianesi, B. and J. V. Reenen (2003). The returns to education: Macroeconomics. *Journal of economic surveys 17*(2), 157–200.

Siedler, T. (2010). Schooling and citizenship in a young democracy: evidence from postwar Germany. *The Scandinavian Journal of Economics 112*(2), 315–338.

Solnick, S. J. and D. Hemenway (2005). Are positional concerns stronger in some domains than in others? *American Economic Review 95*(2), 147–151.

Solnick, S. J., L. Hong, and D. Hemenway (2007). Positional goods in the United States and China. *The Journal of Socio-Economics 36*(4), 537–545.

Spence, A. M. (1973, August). Job market signaling. *Quarterly Journal of Economics 87*(3), 355–374.

Steinerman, A., J. Volshteyn, and M. McGarrett (2011). *Education services data book*. New York: JP Morgan North American Equity Research, Business and Education Services.

Stiglitz, J. E. (1975, June). The theory of 'screening,' education, and the distribution of income. *American Economic Review 65*(3), 283–300.

Stinebrickner, R. and T. R. Stinebrickner (2008). The causal effect of studying on academic performance. *The BE Journal of Economic Analysis & Policy 8*(1), 1–53.

Teulings, C. and M. Koopmanschap (1989). An econometric model of crowding out of lower education levels. *European economic review 33*(8), 1653–1664.

Trostel, P. A. (1993). The effect of taxation on human capital. *Journal of Political Economy 101*(2), 327–350.

Trostel, P. A. (1996). Should education be subsidized? *Public Finance Review 24*(1), 3–24.

Vila, L. E. (2000). The non-monetary benefits of education. *European journal of education 35*(1), 21–32.

Winston, G. and D. Zimmerman (2004). Peer effects in higher education. In *College choices: The economics of where to go, when to go, and how to pay for it*, pp. 395–424. University of Chicago Press.

Winter, C. (2014). Accounting for the changing role of family income in determining college entry. *The Scandinavian Journal of Economics 116*(4), 909–963.

Wolfe, B. L. and R. H. Haveman (2002). Social and nonmarket benefits from education in an advanced economy. In Y. K. Kodrzycki (Ed.), *Education in the 21st Century: Meeting the Challenges of a Changing World*. Boston, MA: Federal Reserve Bank of Boston.

Zarate, M. E. and H. P. Pachon (2006). Perceptions of college financial aid among California Latino youth. Tomas Rivera Policy Institute Policy Brief.

Appendix: proofs

An educational subsidy financed by a poll tax is inequitable

Consider an educational subsidy S paid by lump-sum taxes T. The budget constraint requires

$$T = SP^{\circ},$$

with θ° implicitly defined via

$$b(\theta^{\circ}) + S - c(\theta^{\circ}) = 0. \tag{21}$$

So,

$$\frac{\partial T}{\partial S} = P^{\circ} + S\frac{\partial P^{\circ}}{\partial S} > 0.$$

Equity E is measured by

$$E = \int (w(\theta)-1)u(\theta)f(\theta)d\theta = \int_{<\theta^{\circ}} (w(\theta)-1)u(\theta)f(\theta)d\theta + \int_{\geq\theta^{\circ}} (w(\theta)-1)u(\theta)f(\theta)d\theta,$$

with $u(\theta) = a(\theta) - T$, if $\theta < \theta^{\circ}$ and $u(\theta) = a(\theta) + b(\theta) - T + S - c(\theta)$, if otherwise. Replacing utilities, we can rewrite E as

$$E = \int (w(\theta) - 1)(a(\theta) - T)f(\theta)d\theta + \int_{\geq\theta^{\circ}} (w(\theta) - 1)(b(\theta) + S - c(\theta))f(\theta)d\theta.$$

Using $\bar{w} - 1 = 0$ and equation (21), we obtain

$$\frac{\partial E}{\partial S} = \int_{\geq\theta^{\circ}} (w(\theta) - 1)f(\theta)d\theta = -P^{\circ}(1 - \bar{w}^{\circ}),$$

with $\bar{w}^{\circ} = \int_{\geq\theta^{\circ}} w(\theta)f(\theta)d\theta/P^{\circ}$. Equity decreases when $\frac{\partial E}{\partial S} < 0$, i.e., when $1 - \bar{w}^{\circ} > 0$, as required.

The cross-derivative of participation

If we differentiate equation (9) with respect to S, we get

$$(1-\tau)b'(\theta^\circ)\frac{\partial\theta^\circ}{\partial S}-(1-\varsigma)c'(\theta^\circ)\frac{\partial\theta^\circ}{\partial S}+1=0.$$

Differentiating again with respect to the tax rate τ, we obtain

$$-b'(\theta^\circ)\frac{\partial\theta^\circ}{\partial S}+(1-\tau)b''(\theta^\circ)\frac{\partial\theta^\circ}{\partial S}\frac{\partial\theta^\circ}{\partial\tau}+(1-\tau)b'(\theta^\circ)\frac{\partial^2\theta^\circ}{\partial S\partial\tau}-(1-\varsigma)c''(\theta^\circ)\frac{\partial\theta^\circ}{\partial S}\frac{\partial\theta^\circ}{\partial\tau}-(1-\varsigma)c'(\theta^\circ)\frac{\partial^2\theta^\circ}{\partial S\partial\tau}=0,$$

leading to

$$\frac{\partial^2\theta^\circ}{\partial S\partial\tau}=\frac{[(1-\varsigma)c''(\theta^\circ)-(1-\tau)b''(\theta^\circ)]\frac{\partial\theta^\circ}{\partial S}\frac{\partial\theta^\circ}{\partial\tau}+b'(\theta^\circ)\frac{\partial\theta^\circ}{\partial S}}{(1-\tau)b'(\theta^\circ)-(1-\varsigma)c'(\theta^\circ)}.$$

So, if $b''<0$ and $c''>0$, then the numerator is negative, while the denominator is positive (given $\varsigma,\tau<1$, $b'>0$ and $c'<0$), as required.

Second-best linear taxes and subsidies

Consider a linear earnings tax scheme (with poll tax T and tax rate τ), a linear subsidy scheme (with educational subsidy S and subsidy rate ς), and a graduate tax scheme (a linear rate η on earnings). The results in section 6.1 are derived in the absence of a graduate tax scheme (or, equivalently, setting $\eta=0$), while the results in section 6.2 are derived with a graduate tax scheme. The payoffs are equal to

$$a(\theta)+\alpha,\ a(\theta)+b(\theta)+\beta, \text{ and } c(\theta)+\gamma,$$

with a strictly increasing, b strictly increasing, c strictly decreasing, and α,β,γ i.i.d. multivariate normal with zero means and $\sigma_{\beta,\gamma}=0$.

The expected utility of an individual with type θ in case he chooses to graduate is equal to

$$EU[-T+(1-\tau-\eta)(a(\theta)+b(\theta)+\beta)+\Xi+S-(1-\varsigma)(c(\theta)+\gamma)],$$

with $\Xi=\int_{P^\circ}e(\theta)f(\theta)d\theta$ the aggregate net effect of participation on society, which is assumed to be non-monetary and therefore non-taxable. Assuming a constant absolute risk aversion $\rho>0$, the expected utility of a normally distributed stochast X (with mean

μ_X and variance σ_X^2) is equal to $\mu_X - \frac{\rho}{2}\sigma_X^2$. So, expected utility becomes

$$-T+(1-\tau-\eta)(a(\theta)+b(\theta))+\Xi+S-(1-\varsigma)c(\theta)-\frac{\rho}{2}[(1-\tau-\eta)^2\sigma_\beta^2+(1-\varsigma)^2\sigma_\gamma^2]. \quad (22)$$

Similarly, the expected utility in case of non-graduation equals

$$EU[-T+(1-\tau)(a(\theta)+\alpha)+\Xi] = -T+(1-\tau)a(\theta)+\Xi-\frac{\rho}{2}(1-\tau)^2\sigma_\alpha^2. \quad (23)$$

An individual with type θ chooses to study if the prospect of studying (equation (22) minus Ξ) dominates the prospect of not studying (equation (23) minus Ξ), or if

$$-\eta a(\theta)+(1-\tau-\eta)b(\theta)+S-(1-\varsigma)c(\theta) > \frac{\rho}{2}\Delta\sigma_{\varsigma,\tau,\eta}^2,$$

with the risk difference $\Delta\sigma_{\varsigma,\tau,\eta}^2$ equal to

$$(1-\tau-\eta)^2\sigma_\beta^2+(1-\varsigma)^2\sigma_\gamma^2-(1-\tau)^2\sigma_\alpha^2.$$

The cutoff type θ° is therefore defined as

$$-\eta a(\theta^\circ)+(1-\tau-\eta)b(\theta^\circ)+S-(1-\varsigma)c(\theta^\circ)-\frac{\rho}{2}\Delta\sigma_{\varsigma,\tau,\eta}^2 = 0.$$

Implicit differentiation leads to

$$\frac{\partial\theta^\circ}{\partial\tau} = \frac{b(\theta^\circ)-\rho[(1-\tau-\eta)\sigma_\beta^2-(1-\tau)\sigma_\alpha^2]}{-\eta a'(\theta^\circ)+(1-\tau-\eta)b'(\theta^\circ)-(1-\varsigma)c'(\theta^\circ)}, \quad (24)$$

$$\frac{\partial\theta^\circ}{\partial\eta} = \frac{a(\theta^\circ)+b(\theta^\circ)-\rho(1-\tau-\eta)\sigma_\beta^2}{-\eta a'(\theta^\circ)+(1-\tau-\eta)b'(\theta^\circ)-(1-\varsigma)c'(\theta^\circ)},$$

$$\frac{\partial\theta^\circ}{\partial\varsigma} = -\frac{c(\theta^\circ)+\rho(1-\varsigma)\sigma_\gamma^2}{-\eta a'(\theta^\circ)+(1-\tau-\eta)b'(\theta^\circ)-(1-\varsigma)c'(\theta^\circ)},$$

and

$$\frac{\partial\theta^\circ}{\partial S} = -\frac{1}{-\eta a'(\theta^\circ)+(1-\tau-\eta)b'(\theta^\circ)-(1-\varsigma)c'(\theta^\circ)}.$$

For later use note that

$$\frac{\partial\theta^\circ}{\partial\tau} = -[b(\theta^\circ)-\rho[(1-\tau-\eta)\sigma_\beta^2-(1-\tau)\sigma_\alpha^2]]\frac{\partial\theta^\circ}{\partial S}, \quad (25)$$

$$\frac{\partial\theta^\circ}{\partial\eta} = -[a(\theta^\circ) + b(\theta^\circ) - \rho(1-\tau-\eta)\sigma_\beta^2]\frac{\partial\theta^\circ}{\partial S}, \tag{26}$$

$$\frac{\partial\theta^\circ}{\partial\varsigma} = [c(\theta^\circ) + \rho(1-\varsigma)\sigma_\gamma^2]\frac{\partial\theta^\circ}{\partial S}. \tag{27}$$

The weighted average expected utility in society is equal to

$$\int w(\theta)EU(\theta)f(\theta)d\theta = \int_{<\theta^\circ} w(\theta)\, EU(\theta)f(\theta)\, d\theta + \int_{\geq\theta^\circ} w(\theta)\, EU(\theta)f(\theta)\, d\theta.$$

Replacing $EU(\theta)$ by (22) and (23) for graduates and non-graduates, social welfare becomes

$$-T + \Xi + (1-\tau)\int w(\theta)a(\theta)f(\theta)d\theta - \frac{\rho}{2}(1-\tau)^2\sigma_\alpha^2 +$$
$$\int_{P^\circ} w(\theta)\,[-\eta a(\theta) + (1-\tau-\eta)b(\theta) + S - (1-\varsigma)c(\theta) - \frac{\rho}{2}\Delta\sigma^2_{\varsigma,\tau,\eta}]f(\theta)\, d\theta.$$

The budget constraint of the policy maker requires that the (expected) net tax revenues are equal to zero, so

$$T+\tau\bar{a}+\int_{\geq\theta^\circ}[\tau b(\theta)+\eta(a(\theta)+b(\theta))]f(\theta)d\theta-\int_{\geq\theta^\circ}(S+\varsigma c(\theta))f(\theta)d\theta-\frac{1}{2}d_\tau\tau^2-\frac{1}{2}d_\eta\eta^2-\frac{1}{2}d_\varsigma\varsigma^2 = 0,$$

with the last three terms capturing the distortions caused by the earnings tax rate τ, the graduate tax rate η, and the subsidy rate ς, all expressed in per-capita revenue terms. We can solve for $-T$ and obtain immediately

$$-T = \tau\bar{a}+\int_{\geq\theta^\circ}[\tau b(\theta)+\eta(a(\theta)+b(\theta))]f(\theta)d\theta-\int_{\geq\theta^\circ}(S+\varsigma c(\theta))f(\theta)d\theta-\frac{1}{2}d_\tau\tau^2-\frac{1}{2}d_\eta\eta^2-\frac{1}{2}d_\varsigma\varsigma^2.$$

Replacing $-T$ and using the formula for Ξ, social welfare is equal to

$$\begin{gathered}\int w(\theta)a(\theta)f(\theta)d\theta - \tau\int(w(\theta)-1)a(\theta)f(\theta)d\theta \\ + \int_{\geq\theta^\circ}[e(\theta) + \tau b(\theta) + \eta(a(\theta)+b(\theta)) - S - \varsigma c(\theta)]f(\theta)d\theta - \tfrac{1}{2}d_\tau\tau^2 - \tfrac{1}{2}d_\eta\eta^2 - \tfrac{1}{2}d_\varsigma\varsigma^2 \\ -\tfrac{\rho}{2}(1-\tau)^2\sigma_\alpha^2 + \int_{\geq\theta^\circ} w(\theta)\,[-\eta a(\theta) + (1-\tau-\eta)b(\theta) + S - (1-\varsigma)c(\theta) - \tfrac{\rho}{2}\Delta\sigma^2_{\varsigma,\tau,\eta}]f(\theta)\, d\theta.\end{gathered}$$

The partial derivatives of welfare with respect to τ, η, ς, and S are

$$\frac{\partial W}{\partial \tau} = \int (1-w(\theta))a(\theta)f(\theta)d\theta + \int_{\geq\theta^\circ}(1-w(\theta))b(\theta)f(\theta)d\theta$$
$$-[e(\theta^\circ)+\tau b(\theta^\circ)+\eta(a(\theta^\circ)+b(\theta^\circ))-S-\varsigma c(\theta^\circ)]f(\theta^\circ)\frac{\partial\theta^\circ}{\partial\tau} - d_\tau\tau$$
$$+\rho(1-\tau)\sigma_\alpha^2(1-\int_{\geq\theta^\circ} w(\theta)\, f(\theta)\, d\theta) + \rho(1-\tau-\eta)\sigma_\beta^2\int_{\geq\theta^\circ} w(\theta)\, f(\theta)\, d\theta,$$

$$\frac{\partial W}{\partial \eta} = \int_{\geq\theta^\circ}(1-w(\theta))(a(\theta)+b(\theta))f(\theta)d\theta$$
$$-[e(\theta^\circ)+\tau b(\theta^\circ)+\eta(a(\theta^\circ)+b(\theta^\circ))-S-\varsigma c(\theta^\circ)]f(\theta^\circ)\frac{\partial\theta^\circ}{\partial\eta} - d_\eta\eta$$
$$+\rho(1-\tau-\eta)\sigma_\beta^2\int_{\geq\theta^\circ} w(\theta)\, f(\theta)\, d\theta,$$

$$\frac{\partial W}{\partial \varsigma} = \int_{\geq\theta^\circ}(w(\theta)-1)c(\theta)f(\theta)d\theta - [e(\theta^\circ)+\tau b(\theta^\circ)+\eta(a(\theta^\circ)+b(\theta^\circ))-S-\varsigma c(\theta^\circ)]f(\theta^\circ)\frac{\partial\theta^\circ}{\partial\varsigma}$$
$$-d_\varsigma\varsigma + \rho(1-\varsigma)\sigma_\gamma^2\int_{\geq\theta^\circ} w(\theta)\, f(\theta)\, d\theta,$$

and

$$\frac{\partial W}{\partial S} = -\int_{\geq\theta^\circ}(1-w(\theta))f(\theta)d\theta - [e(\theta^\circ)+\tau b(\theta^\circ)+\eta(a(\theta^\circ)+b(\theta^\circ))-S-\varsigma c(\theta^\circ)]f(\theta^\circ)\frac{\partial\theta^\circ}{\partial S}.$$

Using the first-order condition for S ($\frac{\partial W}{\partial S}=0$) and equations (25)-(27) we can rewrite the partial derivatives for τ, η, and ς at the optimal $S^\star$ as

$$\frac{\partial W}{\partial \tau}\Big|_{S^\star} = \int (1-w(\theta))a(\theta)f(\theta)d\theta + \int_{\geq\theta^\circ}(1-w(\theta))b(\theta)f(\theta)d\theta - P^\circ b(\theta^\circ)(1-\bar{w}^\circ)$$
$$+\rho(1-\tau)(1-P^\circ)\sigma_\alpha^2 + \rho(1-\tau-\eta)P^\circ\sigma_\beta^2 - d_\tau\tau, \tag{28}$$

$$\frac{\partial W}{\partial \eta}\Big|_{S^\star} = \int_{\geq\theta^\circ}(1-w(\theta))(a(\theta)+b(\theta))f(\theta)d\theta - P^\circ(a(\theta^\circ)+b(\theta^\circ))(1-\bar{w}^\circ) + \rho(1-\tau-\eta)P^\circ\sigma_\beta^2 - d_\eta\eta, \tag{29}$$

and

$$\frac{\partial W}{\partial \varsigma}\Big|_{S^\star} = \int_{\geq\theta^\circ}(w(\theta)-1)c(\theta)f(\theta)d\theta + P^\circ c(\theta^\circ)(1-\bar{w}^\circ) + \rho(1-\varsigma)P^\circ\sigma_\gamma^2 - d_\varsigma\varsigma. \tag{30}$$

We finally rewrite these partial derivatives in covariances. We start with equation (28). Let $y(\theta)$ be ex-post gross income (either $a(\theta)$ or $a(\theta)+b(\theta)$ depending on the participation choice). We have

$$\begin{aligned} cov(w,y) &= \int (w(\theta)-\bar{w})(y(\theta)-\bar{y})f(\theta)d\theta = \int (w(\theta)-1)y(\theta)f(\theta)d\theta \\ &= \int (w(\theta)-1)a(\theta)f(\theta)d\theta + \int_{\geq\theta^\circ}(w(\theta)-1)b(\theta)f(\theta)d\theta. \end{aligned} \tag{31}$$

In other words, the first three terms of equation (28), the net redistributive payoff, are equal to

$$\textit{net redistributive payoff} = -cov(w, y) - P^{\circ}b(\theta^{\circ})(1 - \bar{w}^{\circ}).$$

To see that the net redistributive payoff is positive, we can rewrite the first and the second term of $cov(w, y)$ in equation (31) as

$$\int (w(\theta) - 1)a(\theta)f(\theta)d\theta = \int (w(\theta) - \bar{w})(a(\theta) - \bar{a})f(\theta)d\theta = cov(w, a),$$

and

$$\int_{\geq\theta^{\circ}} (w(\theta) - 1)b(\theta)f(\theta)d\theta = P^{\circ}cov^{\circ}(w, b) - P^{\circ}\bar{b}^{\circ}(1 - \bar{w}^{\circ}),$$

with

$$cov^{\circ}(w, b) = \int_{\geq\theta^{\circ}} (w(\theta) - \bar{w}^{\circ})(b(\theta) - \bar{b}^{\circ})f(\theta)d\theta/P^{\circ},$$

$$\bar{w}^{\circ} = \int_{\geq\theta^{\circ}} w(\theta)f(\theta)d\theta/P^{\circ},$$

and

$$\bar{b}^{\circ} = \int_{\geq\theta^{\circ}} b(\theta)f(\theta)d\theta/P^{\circ}.$$

So, the net redistributive payoff is equal to

$$-cov(w, a) - P^{\circ}cov^{\circ}(w, b) + P^{\circ}(\bar{b}^{\circ} - b(\theta^{\circ}))(1 - \bar{w}^{\circ}),$$

which is positive (all covariances are negative and the last term is positive because $\bar{b}^{\circ} > b(\theta^{\circ})$).

Next, to rewrite equation (29), note that

$$\int_{\geq\theta^{\circ}} (w(\theta) - 1)(a(\theta) + b(\theta))f(\theta)d\theta = P^{\circ}cov^{\circ}(w, a + b) - P^{\circ}(a \bar{+} b)^{\circ}(1 - \bar{w}^{\circ}),$$

with

$$cov^{\circ}(w, a + b) = \int_{\geq\theta^{\circ}} (w(\theta) - \bar{w}^{\circ})(a(\theta) + b(\theta) - (a \bar{+} b)^{\circ})f(\theta)d\theta/P^{\circ} < 0,$$

$$\bar{w}^{\circ} = \int_{\geq\theta^{\circ}} w(\theta)f(\theta)d\theta/P^{\circ},$$

and

$$(a \bar{+} b)^\circ = \int_{\geq \theta^\circ} (a(\theta) + b(\theta)) f(\theta) d\theta / P^\circ.$$

So, the first two terms of equation (29), the net redistributive payoff, is equal to

$$net\ redistributive\ payoff\ = -P^\circ cov^\circ(w, a{+}b) + P^\circ[(a \bar{+} b)^\circ - (a(\theta^\circ) + b(\theta^\circ))](1 - \bar{w}^\circ),$$

which is again positive (the covariance is negative and the second term is positive because $(a \bar{+} b)^\circ > a(\theta^\circ) + b(\theta^\circ)$).

Finally, let $z(\theta)$ be ex-post study costs (either 0 or $c(\theta)$ depending on the participation choice). We have

$$cov(w, z) = \int (w(\theta) - \bar{w})(z(\theta) - \bar{z}) f(\theta) d\theta = \int (w(\theta) - 1) z(\theta) f(\theta) d\theta = \int_{\geq \theta^\circ} (w(\theta) - 1) c(\theta) f(\theta) d\theta.$$

In other words, the first two terms of equation (30) are equal to

$$net\ redistributive\ payoff\ = cov(w, z) + P^\circ c(\theta^\circ)(1 - \bar{w}^\circ).$$

To verify the sign again, we can rewrite the covariance as

$$cov(w, z) = \int_{\geq \theta^\circ} (w(\theta) - 1) c(\theta) f(\theta) d\theta = P^\circ cov^\circ(w, c) - P^\circ \bar{c}^\circ (1 - \bar{w}^\circ),$$

with

$$cov^\circ(w, c) = \int_{\geq \theta^\circ} (w(\theta) - \bar{w}^\circ)(c(\theta) - \bar{c}^\circ) f(\theta) d\theta / P^\circ > 0,$$

$\bar{w}^\circ$ as before, and

$$\bar{c}^\circ = \int_{\geq \theta^\circ} c(\theta) f(\theta) d\theta / P^\circ.$$

The net redistributive payoff is now equal to

$$P^\circ cov^\circ(w, c) + P^\circ (c(\theta^\circ) - \bar{c}^\circ)(1 - \bar{w}^\circ),$$

which is again positive (the covariance is positive and the second term is positive because $c(\theta^\circ) > \bar{c}^\circ$).