Designing Efficient College and Tax Policies*

Sebastian Findeisen
University of Mannheim

Dominik Sachs
University of Cologne

November 12, 2014

Abstract

The total social benefits of college education exceed the private benefits because the government receives a share of the monetary returns in the form of income taxes. We study the policy implications of this fiscal externality in an optimal dynamic tax framework. Using a variational approach we derive a formula for the revenue effect of an increase in college education subsidies and an expression for the excess burden of income taxation caused by the college margin. We also study optimal nonlinear income taxation and show that optimal taxes are typically smaller because of the college margin. Our model addresses several important dimensions raised by structural labor economists: multidimensional heterogeneity, idiosyncratic risk and borrowing constraints. The model matches key empirical results on college enrollment patterns, returns to education and enrollment elasticities. Quantitatively, we find that a marginal increase in college subsidies in the US is at least 70 percent self-financing through the net-present value increase in future tax revenue. When targeting this increase to children in the lowest parental income tercile, it is even up to 165 percent self-financing. The excess burden of income taxation is only slightly altered by the college margin and therefore the optimal Mirrleesian income tax schedule is barely affected as well, in particular if subsidies are set at their optimal level.

JEL-**classification:** H21, H23, I22, I24, I28

Keywords: Optimal Taxation, College Enrollment, Tax Reforms

*Contact: findeisen@uni-mannheim.de, sachs@wiso.uni-koeln.de. We are indebted to Aleh Tsyvinski for invaluable guidance, encouragement and support. We are grateful to Felix Bierbrauer, Friedrich Breyer, Bas Jacobs, Alex Ludwig and Emmanuel Saez for helpful discussions and support. We further thank Christian Bredemeier, David Card, Alexander Gelber, Marcel Gerard, Emanuel Hansen, Leo Kaas, Marek Kapicka, Kory Kroft, Fabian Krueger, Paul Klein, Tim Lee, Lance Lochner, Normann Lorenz, Thorsten Louis, Michael Peters, Gianluca Violante, Matthew Weinzierl, Nicolas Werquin and seminar participants at Berkeley, NBER SI Macro Public Finance, Bonn (MPI & MEE), Cologne, Konstanz, Uppsala, SEEK, the Taxation Theory Conference, NORMAC and IIPF. Dominik Sachs' research was partly funded by a post-doc fellowship of the Fritz-Thyssen Foundation and the Cologne Graduate School in Management, Economics and Social Sciences. We gratefully acknowledge the hospitality of Yale and Berkeley where parts of this research were conducted.

1 Introduction

College graduates in the US on average earn $480,000 more over their lifecycle than individuals without a college degree (Autor 2014). According to recent OECD estimates, they contribute $205,003 more to the government budget.[1] These numbers are not directly informative for policy makers because they reflect returns as well as selection, and because returns for marginal students can be quite different. Nevertheless they suggest that income taxation can lift the social returns to education substantially above the private returns, which implies that individual college education decisions impose a *fiscal externality*. In this paper, we theoretically and quantitatively assess the implications of this fiscal externality for the efficient design of college education subsidies and income taxes in the US.

Specifically, we ask: (i) To what degree is an increase in college subsidies in the US self-financing through higher tax revenue in the future? Or, equivalently, what are the actual marginal fiscal costs of college education subsidies? How do these numbers differ for need-based policies? (ii) Does the individual college education margin limit the ability of the government to raise tax revenue because this college margin adds a substantial excess burden? (iii) How is the equity-efficiency trade-off of Mirrlees (1971) affected by the college margin and how are the quantitative implications for optimal nonlinear taxes (Diamond 1998, Saez 2001) altered?

We study a dynamic optimal taxation framework, where individuals decide on college entry early in their life. The framework incorporates heterogeneity in ability, parental income and preferences for college, as well as idiosyncratic uncertainty and borrowing constraints. Our first theoretical contribution is to derive an expression for the revenue effect of an increase in college education subsidies. The direct negative revenue effect increases in the share of inframarginal students, i.e. those individuals that would have enrolled also in the absence of the reform. The positive revenue effect through the fiscal externality increases in the following three measures: the share of marginal students that is induced to enroll because of a small increase in subsidies, the returns to college for these marginal students and the progressivity of the income tax. Our second theoretical contribution is an expression for the excess burden of income taxation caused by the distortion of incentives for college enrollment. For this purpose, we use variational arguments and consider various perturbations of nonlinear income tax schedules.[2] Our third theoretical contribution is to explore optimal policies in the spirit of Mirrlees (1971). We study policy instruments with various degrees of sophistication: optimal subsidies for given income taxes and vice versa, jointly optimal subsidies and income taxes as well as optimal education dependent tax systems.

[1] The estimate from Autor (2014) reflects the difference in the net present value of lifetime earnings net of tuition. The OECD numbers are taken from Table A.74 in OECD (2014).

[2] For the general variational approach to the analysis of tax systems in dynamic environments, see Golosov, Tsyvinski, and Werquin (2014). In a static framework, this method has been applied to characterize the optimal tax schedule by Piketty (1997) and Saez (2001).

Our modeling assumptions are strongly guided by the recent structural labor literature on college education (e.g., Johnson 2013, Abbott, Gallipoli, Meghir, and Violante 2013). Thus, our framework connects theoretical models of dynamic policy reforms and Mirrleesian income taxation with structural models of college education. We estimate the relevant parameters of our model using data of the National Longitudinal Survey of Youth (NLSY). We also provide a detailed modeling of the US income tax schedule and college education policy system. Our quantitative model matches enrollment patterns in terms of ability and parental income. As we argue in detail below, our model very well replicates quasi-experimental results about the responsiveness of the college margin with respect to subsidies and parental income that we do not target in our estimation (Deming and Dynarski 2009, Castleman and Long 2013, Hilger 2014). The implied (marginal) returns to college are in line with the returns found by labor economists (Card 1999, Oreopoulos and Petronijevic 2013, Zimmermann 2014).

The combination of analytical tractability and empirical plausibility of our model allows for a comprehensive quantitative analysis of policy reforms and optimal policies. Our quantitative results are: (i) A marginal increase in college subsidies is 70-87% self-financing through higher discounted future tax revenue. It is a free lunch if the increase is targeted to children in the lowest parental income tercile: for each dollar of subsidies the government obtains between \$1.00 and \$1.65 additional income taxes in net present value. This is mainly driven by the lower share of inframarginal students among those children with low parental income. The lower bounds of the values take into account dropout and adverse effects on parental labor supply. Allowing for general equilibrium effects on the fiscal externality makes the results, perhaps surprisingly, stronger. (ii) The marginal excess burden of income taxation caused by distorted incentives for college education lies below \$0.01 for each tax dollar raised. Thus, for each mechanically raised dollar of tax revenue, the government loses less than 1 Cent through the disincentives for college education. (iii) The optimal Mirrleesian income tax schedule is barely influenced by the endogeneity of the college education decision. Optimal marginal tax rates for given college subsidies are at most 1.2 percentage points lower as compared to the case with exogenous college enrollment decisions. For top incomes, the difference is negligible. If subsidies are set optimally, the difference is negligible for all income levels.

To describe our approach more formally, we consider a model where individuals decide whether to work directly after high school or postpone labor market entry and obtain a college degree.[3] In the latter case, individuals make a borrowing decision during college. In either case, they face idiosyncratic wage risk that materializes after entering the labor market. The distribution from which they draw their wage depends on their education decision. After individuals have learned their wage, they make a labor leisure decision. Labor supply is endogenous with

[3]In the baseline model we abstract from college dropout. We extend our analysis and consider dropout for the quantitative analysis of college subsidies.

respect to the income tax system. Ex-ante, individuals differ with respect to ability, psychic costs of going to college and parental income. Ability determines the wage distributions, and therefore the returns to college as well as the opportunity costs of foregone earnings. Parental income matters because of (potential) borrowing constraints and because parental transfers typically depend on the education choice of the children (besides parental income). The latter effect acts like a price subsidy. Despite the large amount of heterogeneity in the model, we can provide an analytical treatment of optimal education policies and arbitrarily nonlinear and potentially education dependent income taxes. The reason is that we focus on preferences that abstract from income effects on labor supply; this allows us to keep the screening problem one-dimensional once uncertainty has materialized. Intuitively, once the wage is drawn, other variables (apart from income taxes) do not matter for the labor supply decision – a method that we previously explored in Findeisen and Sachs (2014).

To estimate the relevant parameters of the model, we use data from the National Longitudinal Survey of Youth 1979 and 1997 (henceforth, NLSY79 and NLSY97). Parental income and ability are directly taken from the data. For the latter, we use the *Armed Forced Qualification Test Score* (AFQT). Using simple regressions, we estimate how parental income and the AFQT-score determine other variables of our model: we estimate (i) the wage distributions conditional on AFQT and the education decision, (ii) parental transfers conditional on parental income and the education decision, as well as (iii) college grant receipt as a function of ability and parental income. We define each individual that obtained at least a bachelor degree as a college graduate. Based on these estimations, some standard assumptions on preferences and a detailed modeling of the year 2000 college education finance and income tax system,[4] we can calculate the value functions for college and high school graduates for each individual according to our model. Finally, we close the quantitative model by estimating the distribution of psychic costs for college with maximum likelihood to fit college decisions in the data. In line with the empirical literature discussed below, we choose parental education and AFQT as determinants of psychic costs. Our quantitative model can replicate key patterns on how college education varies with parental income and AFQT. Further, our model yields returns to college that are in line with the empirical literature.

We also test the ability of our model to replicate quasi-experimental studies that our estimation did not target. We find that a $1,000 dollar increase in college grants for all students induces an increase in the share of individuals that hold a college degree by 1.48 percentage points. A number that is well in line with the empirical literature surveyed by Kane (2006) and Deming and Dynarski (2009). Further, we find that a $1,000 Dollar increase in parental income triggers a 0.10 percentage points increase in college graduation. A number that is in line – though slightly smaller – with Hilger (2014).

[4]We take the year 2000 because individuals in the NLSY97 are 18 on average in that year.

Our first set of main results concern the impact of a college subsidy increase on future tax revenue. Our starting point is a stylized version of the US college education finance and tax policy system. Our analysis suggests that an increase in subsidies is close to self-financing: a \$1 increase in subsidies could increase discounted future tax revenue by \$0.70 to \$0.87, where the lower number takes into account college dropout. We show that – besides marginal returns to college education – the share of inframarginal and the share of marginal students (or equivalently, the semi-elasticity of enrollment) is key to understand the revenue and welfare implications of an increase in subsidies. As we argue in detail below, our quantitative model replicates empirical evidence about these shares very well. Intuitively, the fiscal gains of higher subsidies are proportional to the share of marginal students (given that they pay more taxes) and the marginal fiscal costs are proportional to the share of inframarginal students (to whom higher subsidies are paid without changing their behavior). *Ceteris paribus* a lower share of inframarginal students makes such a policy reform more desirable. In accordance with this consideration, our analysis suggests that increasing subsidies for children with low parental income provides higher future tax revenue for each marginal dollar invested in subsidies. Although returns to attending college are lower for those income groups, the low share of inframarginal students makes those targeted policies very cost effective. For example, targeting subsidies to children with parental income below the \$35,000 (all dollar values in year 2000 dollars) yields additional tax revenue of \$1.00 to \$1.65 for each invested marginal dollar. Increasing subsidies for these children can thus be seen as an investment in future tax revenue with a rate of return of up to 65%. More generally, our results imply that *efficient* policies target groups with low enrollment rates. Even for a society with no preferences for upward mobility, implementing policies which favor upward mobility are likely to be efficient in this context.

Our second contribution is to measure the excess burden of income taxation through its impact on incentives for college education – in the following often referred to as *the college channel*. To measure the excess burden, we quantify the "leaky bucket" (Okun 1975) and ask how much of each mechanically raised dollar of tax revenue gets lost through the college channel. We consider nonlinear tax reforms and assess the excess burden for increases of marginal tax rates at different income levels; our starting point is a parametric approximation of the US tax system.[5] Our theoretical analysis shows that semi-elasticities of college enrollment with respect to income taxes are a key determinant for this additional excess burden. There is no clear guidance about these semi-elasticities in the literature. We therefore ask how the responsiveness of the college margin to subsidies (for which we have empirical guidance) translates – through the lens of our model – into the responsiveness to income taxes and what this implies for the excess burden. For our benchmark case, we find that at most .6 Cents per dollar are lost

[5]This is in the spirit of Golosov, Tsyvinski, and Werquin (2014), who look at various tax reforms in dynamic models.

through the college channel. This number would result if the marginal tax rate at income level \$70,000 was raised. When increasing marginal tax rates for income levels below \$30,000, an income effect on college education dominates which implies that for each mechanically raised dollar, the government even obtains additional tax revenue by the increase in college graduation. Increasing marginal tax rates for all income levels implies a very small loss of .02 Cents for each mechanically raised Dollar. The responsiveness to subsidies translates into a relatively small responsiveness to taxes because of borrowing constraints and the combination of wage risk and decreasing marginal utility (i.e. risk aversion). We also provide some comparative statics and show that the excess burden can get somewhat larger if borrowing constraints are assumed to be relatively loose and risk aversion is small. In the most extreme case the marginal excess burden caused by the college channel is 1.2 Cents for a uniform marginal tax rate increase.

Finally, we study optimal policies and in particular optimal nonlinear income taxation à la Mirrlees (1971). In line with our results on the excess burden, we find that the structure and size of optimal marginal tax rates is barely influenced by the endogeneity of college enrollment, in particular for top tax rates. Our framework is sufficiently flexible to study optimal taxation in the presence of optimal and suboptimal subsidies. In both cases, optimal Utilitarian income taxes are higher than current taxes; as a result optimal education subsidies are higher than they would be in the presence of the current tax system.[6] To measure the impact of the college channel, our comparison tax schedule is the one that is optimal for a naive social planner that does not take into account the endogeneity of the college decision.[7] The difference between the optimal schedule and that of the naive planner is very small in both cases. It is smaller in the case where subsidies are set optimally because the fiscal externality is – ceteris paribus – smaller for larger subsidies. Finally, we also consider the design of education dependent taxes and find that the excess burden from the college channel is more important than for education independent taxes. In addition, we find that within an optimal education dependent tax system, college graduates face higher marginal tax rates for almost all income levels – a result that is consistent with our findings in Findeisen and Sachs (2013).

This paper is organized as follows. After reviewing the related public finance and empirical college literature in Section 2, we introduce our model in Section 3. In Section 4, we present our empirical estimation of the relevant model parameters and show how our model can match key empirical features. In Section 5, we analytically derive simple formulas for the impact of increases in education subsidies for public funds and provide the respective numbers based on our estimated model. We derive formulas to assess the excess burden of taxation caused by the college channel in Section 6 and also provide a quantitative treatment. In Section 7, we

[6] This follows the "siamese twins" logic of Bovenberg and Jacobs (2005).

[7] Formally, this comparison can be considered as a self-confirming policy equilibrium as defined in Rothschild and Scheuer (2013,14).

derive optimality conditions in the tradition of the Mirrlees taxation literature for policies with various degrees of sophistication and present our quantitative results. Section 8 concludes.

2 Related Literature

2.1 Public Finance

We derive our main theoretical results using intuitive tax perturbation methods, which go back to Piketty (1997) and Saez (2001) in the static context and have recently been generalized to dynamic settings by Golosov, Tsyvinski, and Werquin (2014). As the latter authors, we use the formulas to calculate revenue effects of policy reforms. As the former authors, we also use these formulas to calculate optimal policies. Our paper provides an extension of their results to an environment with endogenous college enrollment decisions.

Concerning the results for optimal income taxes, we show how our formula is an extended version of the well known Diamond (1998) formula. Since college enrollment is modeled as a binary choice, our formal approach is similar to other optimal tax papers with both, intensive and extensive margin. Scheuer (2014) considers the occupational choice margin (become an entrepreneur or not), Saez (2002) and Jacquet, Lehmann, and Van der Linden (2013) consider the labor force participation margin and Lehmann, Simula, and Trannoy (2014) consider migration.

An important theoretical benchmark for a public finance treatment of education is the paper by Bovenberg and Jacobs (2005). They have shown that education subsidies and redistributive taxes are "siamese twins" in that the former can counteract the distortion on the education margin of the latter. Our paper aims at quantifying these effects in the context of college education in the US, for which we need a model that contains idiosyncratic uncertainty and is richer in terms of heterogeneity. In addition, besides optimal policies we also elaborate on the implications of these arguments for small changes of *current* policies.[8]

Bohacek and Kapicka (2008) extend the findings of Bovenberg and Jacobs (2005) to a dynamic setting and provide a quantitative exploration of the theory as well. More recently, Kapicka and Neira (2013), Findeisen and Sachs (2013) and Stantcheva (2014) consider the issue in the "New Dynamic Public Finance" (NPDF) tradition with endogenous policy instruments.[9] Our paper differs in that we explicitly consider college as an education decision and also analyze simpler policy instruments as opposed to solely focusing on the second-best.

[8]Richter (2009) studies a Ramsey setting and finds that qualified labor should be distorted less heavily than non-qualified labor. Further, he shows that the education choice should not be distorted if the earnings function is iso-elastic with respect to educational investment. Anderberg (2009) studies the implications of the risk properties of human capital for optimal human capital policy rules. Jacobs, Schindler, and Yang (2012) consider the implications of earnings risk for optimal linear education subsidies and income taxes.

[9]Kapicka (2013) studies optimal taxation over the life cycle with unobservable human capital and finds a force towards decreasing taxes over the life cycle to incentivize human capital investment. Best and Kleven (2013) make a similar case, connecting optimal tax formulas to empirically estimated career effects.

Krueger and Ludwig (2013a, 2013b) solve for the optimal linear labor income tax rate (with a lump sum rebate) and the optimal college subsidy in a large scale macroeconomic overlapping generations model. We view our analysis as complementary in that we provide analytical results and our policy space is richer whereas they consider a larger model in the macroeconomic tradition. Finally, we have a strong focus on reforms of *current* policies besides characterizing optimal policies.

Hamilton (1987) has argued that taxing capital can be optimal in the presence of labor income taxation. Since capital is a substitute for human capital, a tax on interest income boosts human capital investment, which is beneficial because it counteracts the labor income tax distortion on the human capital margin. Jacobs and Bovenberg (2010) elaborate this argument in a Mirrlees income tax setting. Schindler and Yang (2014) consider whether education subsidies or capital taxes are the better catalyst for social insurance and derive a Ramsey rule for the optimal use of these two instruments. Ábrahám, Koehne, and Pavoni (2014) study an optimal social insurance model in the spirit of Varian (1980), where individuals make an unobservable effort decision in the first period, which can be interpreted as a human capital investment. Capital taxation is desirable in such a setting because it curbs the moral hazard problem. If it is not possible to tax capital, they show that optimal labor income taxes become less progressive. This is related to out results in Section 7, where we quantitatively show that optimal income taxes are less progressive if college can be subsidized optimally.

From a methodological point of view, our paper is related to the work of Blundell and Shephard (2012), who explore optimal income taxation in a structurally estimated model of labor supply. Whereas our empirical approach is less ambitious in that we set some parameters (labor supply elasticity, risk aversion, discount factor) in line with benchmark numbers from the literature instead of estimating them, our approach has the advantage of providing an analytical treatment of the policy problem making our results particularly transparent.

For our results on college subsidies, a close recent paper is Lawson (2014). He uses a sufficient statistic approach for the design of optimal college tuition subsidies. In contrast to his work, we focus on revenue effects of small reforms and we allow for heterogeneity in returns and parental income. This is an important feature, which allows us to analyze targeted subsidies such as need-based or merit-based financial aid.[10]

Badel and Huggett (2014) study how the Laffer curve for the top income tax rate changes if human capital accumulation is taken into account. They show that applying the formula of Diamond and Saez (2011) for the Laffer top tax rate yields values that are are significantly higher than the actual Laffer rate. We find that the formula from static optimal tax theory that ignores the college margin yields values for the top tax rate that barely differ from the

[10]Holter (2014) studies the impact of public educational expenditure and income taxes on intergenerational earnings persistence and as a side result finds that implementing the Danish education subsidies in the US would be self-financing.

optimal values that take the college margin into account. These results are not contrary as they study a structural model with human capital accumulation over the life cycle.

2.2 Empirical College Decision Analysis

Several papers in labor and education economics have described college decisions of individuals and how they respond to changes in policies. Our analysis is strongly guided by this empirical literature. We discuss the most relevant papers in this subsection.

A considerable reduced form literature has investigated the responsiveness of college enrollment with respect to education policies. A benchmark result from quasi-experimental studies for the US is that an increase in yearly college grants of $1,000 leads to an increase in college enrollment by 3-5 percentage points (Kane 2006, Deming and Dynarski 2009).[11] For our quantified model, we arrive at similar numbers when performing such policy experiments. Concerning the impact of parental income/wealth on college enrollment, empirical evidence is less conclusive. See, e.g., Lovenheim (2011) and Hilger (2014) for two recent papers with differing results. Lovenheim (2011) uses variation in housing wealth with the PSID as data source, finding relatively big effects. Hilger (2014) exploits differences in the timing of parental layoffs using IRS micro data and finds small effects: a $1,000 increase in parental income induces an increase in college enrollment of 0.43 percentage points. By increasing parental income by $1,000 for all individuals in our estimated model, we arrive at numbers that are slightly more than half the size of those in Hilger (2014).[12]

Empirical evidence on the importance of borrowing constraints is also less conclusive. Belley and Lochner (2007) find that parental income and wealth have a stronger impact on the enrollment decision than 30 years ago. They set up an educational choice model and find that the model can explain this increasing importance only by borrowing constraints. Johnson (2013) considers a state of the art structural college model with three choices (no college, 2 years college or 4 years college). He finds that relaxing borrowing constraints only has a modest impact on enrollment and that an increase in tuition subsidies is needed in order to obtain a larger increase in college enrollment. Navarro (2011) reaches different conclusions in a semi-parametrically identified structural model. Lochner and Monge-Naranjo (2011) explicitly distinguish between private and public lending and endogenize the privat borrowing limit and can explain more empirical patterns. Winter (2014) uses a structural life-cycle model of college entry where children receive transfers from their altruistic parents and finds that a quarter of

[11] Performing such policy experiments within structural models, Johnson (2013) and Abbott, Gallipoli, Meghir, and Violante (2013) arrive at similar numbers. Using a semi-structural approach, Navarro (2011) arrives at lower numbers. His results suggest that abolishing tuition would increase enrollment by only 2 percentage points.

[12] Very recently, Manoli and Turner (2014) find relatively large cash-on-hands effects exploiting two kink points in the US federal income tax code.

all households are financially constrained. This whole issue is comprehensively surveyed by Lochner and Monge-Naranjo (2012). In our specification of the quantitative model, we assume that individuals can only borrow trough the Stafford loan system at a maximum of $23,000 which was the relevant limit for the time period we consider. We find that a removal of borrowing constraints leads to an increase of 4.2 percentage points in college graduation rates, a number that is a bit larger than that of Johnson (2013), which is 2.4 percentage points.

Besides borrowing constraints, another reason for the strong link between parental income and college enrollment are parental transfers. For various reasons, parents might subsidize the college decision of their children and the extent of such parental subsidies naturally varies with income. We follow Keane and Wolpin (2001) and Johnson (2013) and model parental transfers in a reduced form fashion.

There is also strong empirical evidence that returns to college cannot be the only determinant of the college entry decision. Other variables that are considered in the literature as driving forces are parental education, ability and the geographical environment. Taken together these variables are often summarized by one variable labelled psychic costs or tastes for college, see Cunha, Heckman, and Navarro (2005), Heckman, Lochner, and Todd (2006), Cunha, Karahan, and Soares (2011), Navarro (2011) and Johnson (2013).

Abbott, Gallipoli, Meghir, and Violante (2013) provide a state of the art large scale college enrollment model with endogenous parental transfers and imperfect substitutability between college and high-school labor. Similar as Johnson (2013) they find that increasing loan limits would have no salient effect on enrollment. For grant increases they obtain similar numbers. However, when taking into account general equilibrium effects (i.e. that the college wage premium declines when graduation rates increase), the effects of grant increases on enrollment are only 1/3 or 1/4 as big.

Finally, Johnson (2013) and Abbott, Gallipoli, Meghir, and Violante (2013) find that need based college grants are more cost-effective in increasing enrollment than merit-based grants because among low income individuals, the share of inframarginal students is low. This is related to our result that an increase in need-based financial aid is likely to be a free lunch.

3 The Formal Framework

This section is organized as follows: In Section 3.1, the economic environment is introduced. In Section 3.2, we describe individual behavior and the equilibrium given taxes. In Section 3.3, we derive analytical expressions for behavioral responses along the extensive college margin with respect to various policy changes.

3.1 Individual Heterogeneity, Preferences and Choices

Short Verbal Summary of the Model. Individuals start life as high school graduates and decide whether to obtain a college degree. They differ in innate ability, their preferences for college and parental income. If an individual decides against a college degree, she directly enters the labor market, draws a lifetime wage from a distribution that depends on her innate ability, and – conditional on the drawn wage – makes a labor leisure decision. If an individual goes to college, this takes a certain amount of time and the individual additionally makes a savings decision. Once the individual graduates from college, she enters the labor market, draws a lifetime wage from a different distribution that also depends on her innate ability, and makes a labor-leisure decision.

Whether an individual attends college depends on innate ability through the returns to college and the opportunity costs of foregone earnings. It also depends on parental income, as we allow *parental transfers* to be determined by parental income. First, parental transfers matter because of potentially binding borrowing constraints. Second, parental transfers act in a similar way as a price subsidy because parents make transfers contingent on the educational decision. Lastly, obtaining a college degree has some direct effect on utility which can be positive for some individuals (pleasure from studying) and negative for other individuals (effort costs of studying).

Heterogeneity. Individuals differ in their ability $\theta \in \Theta := [\underline{\theta}, \overline{\theta}]$, parental income $\mathcal{I} \in \mathbb{R}_+$ and an N-dimensional vector of other characteristics X. The joint probability density function (*pdf*) is given by $j(\theta, \mathcal{I}, X)$. Whereas θ and $\mathcal{I}$ are variables that influence an individual's utility through the budget constraint, we assume that X contains variables that influence an individual's decision to go to college by factors outside the budget constraint. All these variables translate into one scalar which we denote *psychic costs* of going to college. Psychic costs κ are thus a function of X. We allow also θ to influence these psychic costs. Hence, we can write $\kappa = z(\theta, X)$.[13] The individual college attendance decision then depends on the triple $(\theta, \mathcal{I}, \kappa)$. The joint density function $m(\theta, \mathcal{I}, \kappa)$ is given by

$$m(\theta, \mathcal{I}, \kappa) = \int_{\{x \in X | k(\theta, x) = \kappa\}} j(\theta, \mathcal{I}, x) dx.$$

We denote the marginal density of ability as $f(\theta)$ (the associated *cdf* is $F(\theta)$), the density of parental income conditional on innate ability as $k(\mathcal{I}|\theta)$ (the associated *cdf* is $K(\mathcal{I}|\theta)$) and the

[13] Our modeling of psychic costs is comparable and close to Cunha, Heckman, and Navarro (2005), Heckman, Lochner, and Todd (2006), Cunha, Karahan, and Soares (2011), Navarro (2011) and Abbott, Gallipoli, Meghir, and Violante (2013), among others.

density of psychic costs conditional on ability and parental income as $h(\kappa|\theta,\mathcal{I})$ (the associated *cdf* is $H(\kappa|\theta,\mathcal{I})$). They are related to the joint density $m(\theta,\mathcal{I},\kappa)$ by

$$m(\theta,\mathcal{I},\kappa) = f(\theta)k(\mathcal{I}|\theta)h(\kappa|\theta,\mathcal{I}).$$

For notational simplicity, we often denote the conditional densities as $h_{\theta,\mathcal{I}}(\kappa)$ and $k_\theta(\mathcal{I})$ and the conditional distribution functions as $H_{\theta,\mathcal{I}}(\kappa)$ and $K_\theta(\mathcal{I})$.

After individuals make the decision to go to college or not, their wage $\omega \in \Omega := [\underline{\omega},\overline{\omega}]$ materializes. If they go to college – i.e. if they are college graduates – they draw their wage from a distribution with *cdf* $G_\theta^{co}(\omega)$ and density $g_\theta^{co}(\omega)$. Otherwise – if they are high-school graduates – they draw their wage from a distribution with *cdf* $G_\theta^{hs}(\omega)$ and density $g_\theta^{hs}(\omega)$. Wage distributions are only influenced by ability and education. In the theoretical part of this paper, we do not make any further assumptions on these wage distributions; note that they encompass labor market returns from innate ability, returns to college, a possible complementarity between ability and schooling as well as the degree of uncertainty.

Preferences. We assume quasi-linear preferences of the form $U\left(c - v\left(\frac{y}{\omega}\right)\right)$, where c denotes consumption, y denotes gross income, $U(\cdot)$ is a concave transformation and $v(\cdot)$, disutility of labor effort, is a strictly convex transformation that satisfies $v(0) = 0$. We make the quasi-linearity assumption for tractability.[14] This assumption allows us to solve for optimal nonlinear income taxes á la Mirrlees (1971) despite the the large amount of heterogeneity. Concretely, the labor supply decision of an individual only depends on ω and policies; it is independent of the amount of non-labor income or wealth. This implies that the decision is independent of X, θ and $\mathcal{I}$.[15] This renders the screening problem one-dimensional. Further, we assume that individuals are expected utility maximizers. Denote by $c_{co}(\omega,\theta,\mathcal{I})$ and $y_{co}(\omega)$ consumption and before tax earnings of a college graduate with wage ω, innate ability θ and parental income $\mathcal{I}$, and by $c_e(\theta,\mathcal{I})$ the respective consumption during college.[16] Expected utility of a college graduate with innate ability θ and parental income $\mathcal{I}$ is given by

$$\beta^{co1}U(c_e(\theta,\mathcal{I})) - \kappa + \beta^{co2}\int_\Omega U\left(c_{co}(\omega,\theta,\mathcal{I}) - v\left(\frac{y_{co}(\omega)}{\omega}\right)\right)dG_\theta^{co}(\omega),$$

[14]The empirical literature using micro data sets has typically found only small income effects on labor supply (see Gruber and Saez (2002) for the US or a recent paper by Kleven and Schultz (2012) using the universe of Danish tax records). In macroeconomics, this class of preferences has shown to be very useful in matching business cycle moments (Greenwood, Hercowitz, and Huffman 1988, Mendoza and Yue 2012).

[15]In Findeisen and Sachs (2014), we elaborate this idea to study simple income taxation in a dynamic stochastic environment, i.e. we characterize income taxes that are simple in the sense that they are based on current annual income as opposed to the whole history of incomes.

[16]Consumption decisions are independent of κ because psychic costs enter the utility function in a separable way. As explained above, earnings decisions are independent of θ and $\mathcal{I}$ because we abstract form income effects.

where β^{co1} and β^{co2} capture discounting and the different length of periods. Assume that life after the college entry decision lasts T periods, college takes T_e periods and individuals' yearly discount factor is β. Then we can think of $\beta^{co1} = \sum_{t=1}^{T_e} \beta^{t-1}$ and $\beta^{co2} = \sum_{t=T_e+1}^{T} \beta^{t-1}$. As is a standard assumption in the literature on college enrollment, psychic costs κ and consumption enter utility separably. For simplicity and without loss of generality, we assume that psychic costs do not accrue every period in college but only once.[17] Note that κ can also take negative values, which then reflects utility benefits from attending college. κ should be interpreted as a one-dimensional aggregate of different factors (i.e. the vector X) that determine the college entry decision by factors outside the budget constraint. In the empirical section of this paper, we will be more explicit about these factors.

Denote by $c_{hs}(\omega, \mathcal{I})$ consumption of a high school graduate with wage ω and parental income $\mathcal{I}$. Note that consumption is not a function of θ in this case, because high school graduates do not make a savings decision – their life is static. $y_{hs}(\omega)$ are before tax earnings. Expected utility of a high school graduate with innate ability θ and parental income $\mathcal{I}$ is then given by

$$\beta^{hs} \int_{\Omega} U\left(c_{hs}(\omega, \mathcal{I}) - v\left(\frac{y_{hs}(\omega)}{\omega}\right)\right) dG_{\theta}^{hs}(\omega),$$

where $\beta^{hs} = \beta^{co1} + \beta^{co2}$ captures the length of the labor market period of high school graduates. Given that life of a high school graduate captures no dynamics, it might seem unnecessary to introduce β^{hs}. However, for the decision of going to college or not, it is important to capture the time aspect to address foregone earnings as opportunity costs of college.

Choices. Individuals first decide on whether to go to college. If individuals attend college, they also make a savings decision. Besides the psychic costs (or benefits) associated with going to college and the foregone earnings, individuals also have to pay tuition fees C. Once they have entered the labor market, individuals make a labor-leisure decision conditional on their realized wage. We do not model college dropout, but show how our results extend to this case in Section 5. While parents are not explicitly modeled, we capture family transfers in a reduced-form manner. Transfers vary with parental income and also with the education decision. We denote the transfers functions by $tr_{hs}(\mathcal{I})$ and $tr_{co}(\mathcal{I})$.[18] In the following subsection, we formally state the individuals' problems and the associated value functions for some given income taxes and education subsidies.

[17] The assumption is harmless because we abstract from college dropout. In Section 5.3, however, we elaborate the robustness of our quantitative findings with respect to dropout.

[18] In general, transfers could also be a function of θ. Our empirical results in Section 4.4 do not support a significant impact of θ on transfers though.

3.2 Equilibrium Given Policies

We next introduce taxes and education subsidies into the model. The tax functions that we consider are potentially education dependent. We denote by $\mathcal{T}^{hs}(y)$ the tax schedule for high school graduates and by $\mathcal{T}^{co}(y)$ the tax schedule for college graduates. $\mathcal{G}(\theta, \mathcal{I})$ is the subsidy or grant that college students receive; the dependence on θ and $\mathcal{I}$ reflects merit and need-based financial aid.

For an individual of type $(\theta, \mathcal{I})$, the value function of being a high school graduate is given by

$$V_{hs}(\theta, \mathcal{I}) = \max_{y^{hs}(w)} \int_{\underline{w}}^{\overline{w}} \beta^{hs}\, U\left(y^{hs}(w) + \frac{tr_{hs}(\mathcal{I})}{\sum_{t=1}^{T} \frac{1}{(1+r)^{t-1}}} - \mathcal{T}_{hs}(y^{hs}(w)) - v\left(\frac{y^{hs}(w)}{w}\right)\right) dG_{\theta}^{hs}(w). \tag{1}$$

A high school graduate essentially solves a static problem as there is no more risk after the initial wage draw. The period length has to be taken into account when transforming the parental transfer into consumption. In the following, we assume $\beta(1+r) = 1$, which implies $\frac{tr_{hs}(\mathcal{I})}{\sum_{t=1}^{T} \frac{1}{(1+r)^{t-1}}} = \beta^{hs} tr_{hs}(\mathcal{I})$. For the college graduate, the problem is slightly more involved:

$$V_{co}(\theta, \mathcal{I}) = \max_{c_e, c^{co}(w), y^{co}(w)} \beta^{co1} U(c_e) + \beta^{co2} \int_{w} U\left(c^{co}(w) - v\left(\frac{y^{co}(w)}{w}\right)\right) dG_{\theta}^{co}(w) \tag{2}$$

subject to

$$\forall\, w: \ \beta^{co2} c^{co}(w) \leq \beta^{co2}\left[y^{co}(w) - \mathcal{T}^{co}(y^{co}(w))\right] + \beta^{co1} s$$

and

$$c_e = \frac{tr_{co}(\mathcal{I})}{\beta^{co1}} + \mathcal{G}(\theta, \mathcal{I}) - C - s,$$

and

$$s \geq \bar{s}.$$

$-s$ is the amount individuals borrow during education. The last constraint implies that individuals are borrowing constrained.[19]

We now turn to the college entry decision. An individual of type $(\theta, \mathcal{I}, \kappa)$ goes to college whenever

$$V_{co}(\theta, \mathcal{I}) - \kappa \geq V_{hs}(\theta, \mathcal{I}).$$

[19] One could easily generalize the approach and assume that the borrowing limit depends on θ or $\mathcal{I}$. Given that we consider $\bar{s}$ being independent of these variables in our quantitative implementation of the model, we stick to this case already at this point of the analysis.

From this expression it follows that there exists a unique threshold of κ for each ability level that divides individuals into college and high school graduates. Formally, this threshold is defined by

$$\tilde{\kappa}(\theta, \mathcal{I}) = V_{co}(\theta, \mathcal{I}) - V_{hs}(\theta, \mathcal{I}) \tag{3}$$

for each $(\theta, \mathcal{I})$.

To ease notation, we define cross sectional wage distributions. First, we define cross sectional distributions conditional on education:

$$G^{co}(\omega) = \int_{\Theta} \int_{\mathbb{R}_+} G^{co}_{\theta}(\omega) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_{\theta}(\mathcal{I}) dF(\theta) \tag{4}$$

and

$$G^{hs}(\omega) = \int_{\Theta} \int_{\mathbb{R}_+} G^{hs}_{\theta}(\omega) \left(1 - H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I}))\right) dK_{\theta}(\mathcal{I}) dF(\theta).$$

Note that G^{co} and G^{hs} are not classical *cumulative distribution functions* because $G^{co}(\overline{\omega}) < 1$ and $G^{hs}(\overline{\omega}) < 1$. Finally, the overall wage distribution is given by

$$G(\omega) = G^{co}(\omega) + G^{hs}(\omega).$$

G is a classical *cdf* because we have $G(\overline{\omega}) = 1$. Note that these wage distributions are endogenous with respect to the tax system. For ease of notation we suppress dependence of these functions with respect to policies.

3.3 Responses Along the College Margin to Policy Changes

In this subsection, we elaborate how the individual college decision responds to changes in policies. Modeling – and quantitatively assessing – behavioral responses along the college margin and assessing the policy implications is the central contribution of this paper.

In general, of course, individuals can respond to policies along three margins: the labor supply margin, the savings margin and the education margin. As individual preferences satisfy quasi-linearity, labor supply will not respond to changes in the absolute value of taxes (and therefore not to changes in the lump sum elements of the tax schedules), but only to changes in marginal tax rates. These responses and their implications for policy design are standard considerations in public economics and well understood.[20] Therefore, we refrain from analyzing labor supply responses in detail. We also do not discuss the responses of savings with respect to policies because we do not consider loan subsidies and related policy instruments.

[20] See Piketty and Saez (2013) for a survey.

3.3.1 Definition of Graduation Semi-Elasticities

College enrollment is responsive to changes in subsidies and the absolute value of taxes. We define *enrollment semi-elasticities* with respect to subsidies and taxes and start with the former:

Grants & Subsidies

$$\xi^{\mathcal{G}}(\theta,\mathcal{I}) \equiv \frac{\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{G}} h_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))k(\mathcal{I}|\theta)f(\theta)}{H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))k(\mathcal{I}|\theta)f(\theta)} = \frac{\beta^{co1}U'(c_e(\theta,\mathcal{I}))h_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))}{H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))} > 0.$$

$\xi^{\mathcal{G}}(\theta,\mathcal{I})$ says by what percentage the share of students of type $(\theta,\mathcal{I})$ increases due to a \$1 increase in grants. The numerator captures the mass of marginal individuals that are induced to go to college due to the subsidy increase. To obtain a semi-elasticity, the mass of those marginal individuals of type $(\theta,\mathcal{I})$ is divided by the share of inframarginal students of type $(\theta,\mathcal{I})$. The equality follows from the fact that $\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{G}} = \beta^{co1}U'(c_e(\theta,\mathcal{I}))$ which one obtains from differentiation of (3). We define semi-elasticities in such a way that a positive semi-elasticity implies an increase in the college graduation rate.

Finally note that semi-elasticities tend to be higher if individuals are borrowing constrained. Intuitively, consumption during college is lower in the presence of credit constraints, which increases the value of $U'(c_e(\theta,I))$.

Taxes We start with the case where taxes are only changed for one education group and first consider taxes for college graduates. We define the semi-elasticity of enrollment for type $(\theta,\mathcal{I})$ with respect to the absolute value of the college tax $\mathcal{T}^{co}(y_{co}(\omega))$ as:

$$\xi^{co}(\theta,\mathcal{I},\omega) \equiv \frac{\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{T}^{co}(y_{co}(\omega))} h_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))k(\mathcal{I}|\theta)f(\theta)}{H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))k(\mathcal{I}|\theta)f(\theta)} = \frac{-\beta^{co2}U'_{co}(\theta,\mathcal{I},\omega)g^{co}_{\theta}(w)h_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))}{H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))} < 0,$$

where $U'_{co}(\theta,\mathcal{I},\omega)$ is the marginal utility of income of a college graduate of type $(\theta,\mathcal{I})$ with realized wage ω. The equality again follows from differentiation of (3). Next, we define the semi-elasticities of enrollment with respect to high school taxes: Marginally increasing $\mathcal{T}^{hs}(y_{hs}(\omega))$ induces a percentage increase in college enrollment for individuals of type $(\theta,\mathcal{I})$ by

$$\xi^{hs}(\theta,\mathcal{I},\omega) \equiv \frac{\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{T}^{hs}(y_{hs}(\omega))} h_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))k(\mathcal{I}|\theta)f(\theta)}{H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))k(\mathcal{I}|\theta)f(\theta)} = \frac{\beta^{hs}U'_{hs}(\mathcal{I},\omega)g^{hs}_{\theta}(\omega)h_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))}{H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))} > 0,$$

where $U'_{hs}(\mathcal{I},\omega)$ is the marginal utility of income of a high school graduate with parental income $\mathcal{I}$ and wage ω. An increase in high school taxes makes college more attractive and increases enrollment: $\xi^{hs}(\theta,\mathcal{I},\omega) > 0$.

Whereas the impact of education dependent tax changes on enrollment is unambiguous, the response of enrollment to an increase of *education-independent* taxes is ambiguous. A marginal increase in $\mathcal{T}(y(\omega))$ yields a percentage increase in enrollment for type θ-individuals given by

$$\xi^{ind}(\theta, \mathcal{I}, \omega) = \xi^{co}(\theta, \mathcal{I}, \omega) + \xi^{hs}(\theta, \mathcal{I}, \omega),$$

which can in general be positive or negative. The next subsection is devoted to determining the sign of $\xi^{ind}(\theta, \mathcal{I}, \omega)$.

3.3.2 Do Higher Taxes Lead to Lower Enrollment?

Let us think about an increase in the absolute tax payment for all individuals with $\omega > \omega^*$ within an education independent tax system, i.e. $\mathcal{T}(\cdot) = \mathcal{T}^{co}(\cdot) = \mathcal{T}^{hs}(\cdot)$. This can be thought of an increase of the marginal tax rate at income level $y(\omega^*)$. Formally, we know from Section 3.3.1 that the overall impact on enrollment reads as

$$\int_{\omega^*}^{\overline{\omega}} \int_{\mathbb{R}_+} \int_{\Theta} \left(\xi^{co}(\theta, \mathcal{I}, \omega) + \xi^{hs}(\theta, \mathcal{I}, \omega)\right) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_\theta(\mathcal{I}) dF(\theta) d\omega.$$

To gain a better understanding, we substitute for the semi-elasticities, yielding:

$$\int_{\omega^*}^{\overline{\omega}} \int_{\mathbb{R}_+} \int_{\Theta} \left(\beta^{hs} U'_{hs}(\mathcal{I}, \omega) g_\theta^{hs}(\omega) - \beta^{co2} U'_{co}(\theta, \mathcal{I}, \omega) g_\theta^{co}(\omega)\right) h_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_\theta(\mathcal{I}) dF(\theta) d\omega. \quad (5)$$

On the one hand, higher income taxes make college graduation more attractive since the tax burden for high school graduates increases (captured by $+\beta^{hs} U'_{hs}(\omega) g_\theta^{hs}(\omega)$), tending to make (5) positive. On the other hand, higher income taxes also affect the individual as a college graduate (captured by $-\beta^{co2} U'_{co}(\omega) g_\theta^{co}(\omega)$), tending to make (5) negative. The question is whether individuals are affected more heavily by this tax increase (in expectation) as a college or as a high school graduate. To understand this, it is useful to distinguish between a price, an income and a time effect.

1. **Price Effect:** This price effect could also be called a return effect. Whenever $G_\theta^{co}(\omega^*) < G_\theta^{hs}(\omega^*)$ (college induces a first-order stochastic dominance shift in the wage distribution), an individual is more likely to be affected by the tax increase if she goes to college. By this reasoning a tax increase for individuals with $\omega > \omega^*$ renders college less attractive.

2. **Income Effect:** Whereas one is more likely to be affected by the tax increase as a college graduate, it is not clear whether one is going to be affected more severely. The "severity" is measured by $E_{\omega|\theta}\left(U'_{hs} | \omega > \omega^*\right)$ and $E_{\omega|\theta}\left(U'_{co} | \omega > \omega^*\right)$. Also here, we have to distinguish two effects.

a) **Wealth Effect:** Whenever college graduates choose to borrow during college, they are poorer as workers than high school graduates for a given realization of ω. This wealth effect implies that higher taxes lead to lower college enrollment and therefore works in the same direction as the price effect. The wealth effect leads to a higher value of $E_{\omega|\theta}\left(U'_{co}|\omega > \omega^*\right)$ relative to $E_{\omega|\theta}\left(U'_{hs}|\omega > \omega^*\right)$.

b) **Expected Wage Effect:** Whereas college graduates are poorer conditional on ω, college graduates – under mild assumptions – have a higher expected wage conditional on $\omega > \omega^*$. Thus, the expected wage conditional on being affected should be lower for a high school graduate than for a college graduate. The expected wage effect leads to a lower value of $E_{\omega|\theta}\left(U'_{co}|\omega > \omega^*\right)$ relative to $E_{\omega|\theta}\left(U'_{hs}|\omega > \omega^*\right)$. Thus, in contrast to the two mentioned effects, the expected wage effect implies that higher taxes lead to higher enrollment.

3. **Time Effect:** A college graduate spends fewer years on the labor market and is affected for a lower number of years by a tax increase. As the expected wage effect, the time effect implies that higher taxes lead to higher enrollment.

We can conclude that the impact of taxes on enrollment is ambiguous. We now consider special conditions under which higher taxes indeed lead to higher enrollment. To achieve this, the price and the wealth effect have to be shut down. The latter is shut down for zero savings and changes its sign for positive savings. The price effect is shut down when we consider increases in lump sum taxes because then everyone is affected with 100% probability.

Lemma 1. *Assume that taxes are education independent, individuals cannot borrow during college and college graduation induces a first-order stochastic dominance shift in wages. Then an increase in the lump sum tax for all individuals unambiguously increases college enrollment.*

Proof. Non-negative savings for college students imply $U'_{hs}(\omega) \geq U'_{co}(\omega) \equiv U'(\omega)$. For a strict equality, we have

$$\int_{\omega} \underbrace{U'_{co}(\theta, \mathcal{I}, \omega)}_{=U'_{hs}(\mathcal{I},\omega)} \left(\beta^{hs} g^{hs}_{\theta}(\omega) - \beta^{co} g^{co}_{\theta}(\omega)\right) d\omega > 0$$

since marginal utility is decreasing in ω. For positive savings, we have $U'_{hs}(\mathcal{I}, \omega) > U'_{co}(\theta, \mathcal{I}, \omega)$ and the argument goes through a fortiori. □

3.3.3 Responsiveness to Taxes versus Subsidies

In the previous subsection, we have argued that higher taxes can increase college graduation. The basic intuition is that higher taxes do not only affect consumption of a college graduate but also of a high school graduate. In this subsection, we look at two policy changes that

have unambiguous effects on graduation and compare the strength of these two effects: grant increases or decreases of taxes for college graduates only.

Lemma 2. Grants versus taxes. *Assume that the borrowing constraint for an individual is binding. Then, an increase in $\mathcal{G}$ by Δ/β^{co1} has a stronger impact on college enrollment than a decrease in $T^{co}(0)$ by Δ/β^{co2} where Δ is a very small number.*

Proof. Differentiation of (3) reveals that $\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{G}} \cdot \Delta/\beta^{co1} = U'(c_e(\theta,\mathcal{I}))\Delta$ and $\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{T}^{co}(0)} \cdot \Delta/\beta^{co2} = \int_\Omega U'(\theta,\mathcal{I},\omega)dG^{hs}_\theta(\omega)\Delta$. If individuals are borrowing constrained, we have $\int_\Omega U'(\theta,\mathcal{I},\omega)dG^{hs}_\theta(\omega) < U'(c_e(\theta,\mathcal{I}))$ and therefore $\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{G}} \cdot \Delta/\beta^{co1} > \frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{T}^{co}(0)} \cdot \Delta/\beta^{co2}$. □

The lemma is straightforward. By construction, a grant increase by $1/\beta^{co1}$ and a lump sum tax decrease for college graduates of $1/\beta^{co2}$ would have the same effect on individual behavior if the individual were not borrowing constrained because they would impact the present value budget constraint in the same manner. If the individual is borrowing constrained, a grant increase with the same net present value is more valuable since it relaxes the borrowing constraint. We next compare the progressive tax reforms to lump-sum reforms.

Lemma 3. Lump-sum tax reforms versus progressive tax reforms. *A small decrease in $\mathcal{T}^{co}(0)$ by Δ has a stronger impact on college graduation than a decrease in $\mathcal{T}^{co}(y)\ \forall\ y \geq y(\omega^*)$ by $\frac{\Delta}{G^{co}_\theta(\omega^*)}$.*

Proof. Differentiation of (3) reveals that $\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{T}^{co}(0)} \cdot \Delta/\beta^{co2} = \int_\Omega U'(\theta,\mathcal{I},\omega)dG^{hs}_\theta(\omega)\Delta$ and $\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{T}^{co}(y)}|_{y\geq y(\omega^*)} \cdot \Delta/G^{co}_\theta(\omega^*) = \int_{\omega^*}^{\overline{\omega}} U'(\theta,\mathcal{I},\omega)dG^{hs}_\theta(\omega)\Delta/G^{co}_\theta(\omega^*)$. Since individuals are risk averse, we know that $\int_{\omega^*}^{\overline{\omega}} U'(\theta,\mathcal{I},\omega)dG^{hs}_\theta(\omega)/G^{co}_\theta(\omega^*) < \int_\Omega U'(\theta,\mathcal{I},\omega)dG^{hs}_\theta(\omega)$ and therefore $\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{T}^{co}(0)} \cdot \Delta > \frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{T}^{co}(y)}|_{y\geq y(\omega^*)} \cdot \Delta/G^{co}_\theta(\omega^*)$. □

By construction, both policy reforms cost the government the same amount of money in net present value terms. Since individuals are risk-averse, the policy that provides more money in bad states of the world is more effective in increasing graduation. This is an interesting result that indicates that marginal tax increases for high incomes are less distortive in terms of college graduation than marginal tax increases for low incomes. This argument will be elaborated upon in Section 6.1.

4 Quantifying the Model

Section 4.1 contains a description of the data that we use and provides a short explanation of the following procedure. In Section 4.2, we describe our choice of those parameters that we do not calibrate/estimate with data. In Sections 4.3-4.6, we describe our empirical procedure to quantify the remaining distributions and parameters. To provide an overview, Table 1 contains

Table 1: Quantification of the Model

Variable/Function	Quantification	Description of Procedure/Reason
θ	AFQT-score from NLSY97	4.1
$F(\theta)$	Distribution from NLSY97	4.1
$\mathcal{I}$	Parental Income from NLSY97	4.1
$K_\theta(\mathcal{I})$	Distribution from NLSY97	4.1
ω	Calibration from gross income as in Saez (2001)	4.3
$G_\theta^{hs}(\omega)$	Income Regression + Wage Calibration	4.3
$G_\theta^{co}(\omega)$	Income Regression + Wage Calibration	4.3
$tr_{hs}(\mathcal{I})$	Regression	4.4
$tr_{co}(\mathcal{I})$	Regression	4.4
κ	Obtain implied (expected) κ from MLE	4.6
$H_{\theta,\mathcal{I}}(\kappa)$	Approximation of implied κ-distribution from NLSY97	4.6
	Utility Function: $\frac{\left(C-\frac{l^{1+\varepsilon}}{1+\varepsilon}\right)^{1-\gamma}}{1-\gamma}$	
ε	0.5	4.2
γ	2	4.2
	Policies	
$\bar{s}$	Stafford Loan Maximum	4.2
$\mathcal{T}(y)$	Gouveia-Strauss Approximation from Guner et al. (2013)	4.2
$\mathcal{G}(\theta,\mathcal{I})$	Two-Step Regression	4.5

a list of all variables and parameters. Finally, in Section 4.7 we evaluate the performance of the model and show that it can replicate key empirical phenomena that have been obtained by quasi-experimental studies.

4.1 Data & Procedure

We use two data sets to bring our model to the data: the National Longitudinal Survey of Youth 79 and 97 (henceforth NLSY79 and the NLSY97). A big advantage of these data sets, which has been exploited in many previous papers, is that they contain the *Armed Forced Qualification Test Score* (AFQT-score) for most individuals, which is a cognitive ability score for high school students that is conducted by the US army. The test score is a good signal for innate ability. Cunha, Karahan, and Soares (2011), e.g., show that it is the most precise signal for innate ability among comparable scores in other data sets.

The NLSY97 is the baseline for our analysis. To quantify the joint distribution of parental income and ability, we take the cross sectional joint distribution in our sample. We then estimate how these variable map into the other variables (parental transfers, wages, grants, psychic costs) of the model. Since individuals in the NLSY97 set are born between 1980 and 1984, not enough information about their earnings is available to quantify the conditional wage distributions. To obtain these conditional wage distributions, we therefore use the NLSY79

data as this data set contains more information about labor market outcomes – individuals are born between 1957 and 1964. The underlying assumption is that the relation between AFQT and wages has not changed over a long time period. Combining both data sets in such a way has proven to be a fruitful way in the literature to overcome the limitations of each individual data set, see Johnson (2013) and Abbott, Gallipoli, Meghir, and Violante (2013). We use the method of Altonji, Bharadwaj, and Lange (2011) to make the AFQT-scores comparable between the two samples and different age groups.

Finally, we define an individual as a college graduate if she has completed at least a bachelor's degree. Otherwise she counts as a high school graduate. Since individuals in the NLSY97 turn 18 years old between 1998 and 2002, we express all US-Dollar amounts in year 2000 dollars.

To quantify our model, we proceed as follows:

1. We make assumptions about some parameters and policies in place in Section 4.2.
2. We estimate $G_\theta^{co}(\omega)$ and $G_\theta^{hs}(\omega)$ in Section 4.3.
3. We estimate $tr_{co}(\mathcal{I})$ and $tr_{hs}(\mathcal{I})$ in Section 4.4.
4. We estimate $\mathcal{G}(\theta, \mathcal{I})$ in Section 4.5.
5. Based on that, we calculate $V_{hs}(\theta, \mathcal{I})$ and $V_{co}(\theta, \mathcal{I})$ for each individual and estimate the distribution of psychic costs with maximum likelihood in Section 4.6.

4.2 Preset Parameters and Policies

We assume the following utility function

$$U = \frac{\left(C - \frac{l^{1+\varepsilon}}{1+\varepsilon}\right)^{1-\gamma}}{1-\gamma},$$

where we set $\varepsilon = 2$, which implies a compensated labor supply elasticity of .5, and $\gamma = 2$. Micro-evidence suggests that the compensated elasticity is probably lower, around .33 (Chetty, Guren, Manoli, and Weber 2011). Given that our elasticity reflects the labor supply responsiveness over the life cycle, we take a larger value of .5. We assume that college takes 4.5 years (i.e. $T_e = 4.5$) and assume that individuals spend 43.5 or 48 years on the labor market depending on whether they went to college before. The choice of 4.5 years for degree completion corresponds to the average years to graduation we observe in the NLSY97, which is 4.57 years. This lines up well with numbers from other sources, for example, from the National Center for Education Statistics (NCES).[21] We set the risk free interest rate to 3%, i.e. $R = 1.03$ and assume that individuals' discount factor is $\beta = \frac{1}{R}$.

[21] See `http://nces.ed.gov/fastfacts/display.asp?id=569`.

To capture current tax policies, we use an approximation of Guner, Kaygusuz, and Ventura (2013) for effective marginal tax rates in the year 2000.[22] We use the year 2000 because individuals in the NLSY97 are 18 in the year 2000 on average. Marginal tax rates concavely increase and converge to roughly 32%. We set the lump sum element of the tax code $\mathcal{T}(0)$ to minus \$1,800 a year. For average incomes this fits the deduction in the US-tax code quite well.[23] For low incomes this reflects that individuals might receive transfers such as food stamps.[24] We set the value of exogenous government spending to 11.2% of the GDP, which is the value that leads to a balanced government budget. This value is a bit low, but this should not be too surprising as we do not take into account corporate taxes or capital income taxes.

For tuition costs, we take average values for the year 2000 from Snyder and Hoffman (2001) for the regions Northeast, North Central, South and West, as they are defined in the NLSY. For all these regions we also take into account the amount of money coming from the taxpayer that is spent per student, which has to be taken into account for the fiscal externality. Both procedures are described in detail in Appendix A.2. The average values are \$7,434 for annual tuition and \$4,157 for the annual subsidy per student. Besides these implicit subsidies, student receive explicit subsidies in the form of grants and tuition waivers. We estimate how this grant receipt varies with parental income and ability in Section 4.5 using information provided in the NLSY. Finally, we make the assumption that individuals can only borrow through the public loan system. In the year 2000, the maximum amount for Stafford loans per student was \$23,000. The latter assumption does not seem innocuous. For our results about the desirability of increasing college subsidies, it is rather harmless because we show how our results can be understood in terms of sufficient-statistics and our quantified model predicts values for these sufficient statistics that are in line with empirical evidence. For the results about the excess burden of income taxation because of the reduced incentives for college education, this assumption plays a role and we provide robustness checks below.

4.3 Estimation of Wage Functions

As a first step, we estimate $G_\theta^{co}(w)$ and $G_\theta^{hs}(w)$. Our approach to estimate the relationship between innate ability, education and labor market outcomes relates to Abbott, Gallipoli, Meghir, and Violante (2013) and Johnson (2013).

[22] We use the "Gouveia-Strauss"-specification including local sales taxes and take the average over all individuals. The parameters can be found in Table 12 of Guner, Kaygusuz, and Ventura (2013).

[23] Guner, Kaygusuz, and Ventura (2013) report a standard deduction of \$7,350 for couples that file jointly. For an average tax rate of 25% this deduction could be interpreted as a lump sum transfer of slightly more than \$1,800.

[24] The average amount of food stamps per eligible person was \$72 per month in the year 2000. Assuming a two person household gives roughly \$1,800 per year. Source: http://www.fns.usda.gov/sites/default/files/pd/SNAPsummary.pdf

In our model, y refers to an average income over the lifetime as we only have one labor market period. Assume that for a high school graduate i, one observes y_t^i for $t = 1, ..., 48$ – i.e. from 18 to 65. The discounted present value of earnings (at age 18) is then given by $\sum_{t=0}^{48} \frac{y_t^i}{(1+r)^{t-1}}$. Simply taking the average over y_t to obtain the relevant income for our model would be misleading since discounting is not taken into account. Thus, we use annuitized income $\tilde{y}_i$ which is given by:

$$\tilde{y}_i = \frac{\sum_{t=1}^{48} \frac{y_t^i}{(1+r)^{t-1}}}{\sum_{t=1}^{48} \frac{1}{(1+r)^{t-1}}}.$$

For some more details on calculating $\tilde{y}_i$, see Appendix A.1. As high school graduate we define everyone with less than 16 years of schooling and as college graduates individuals with 16 or more years of schooling.[25]

The first goal of our empirical strategy aims at getting a mapping between labor earnings, innate abilities, education levels, and their interaction. Comparable to Abbott, Gallipoli, Meghir, and Violante (2013) and Johnson (2013), we run separate regressions, one for high school graduates and one for college graduates, of the form:

$$\ln \tilde{y}_i = \alpha_{ce} + \beta_e^{IN} ln(AFQT_i) + \varepsilon_{ei}^{inc}, \tag{6}$$

for $e = hs, co$. α_{ce} is a cohort-education fixed effect. We find that a one percent increase in AFQT-test scores leads to a 1.88% increase in income for college graduates and 1.28% increase in income for high school graduates, which reflects a complementarity between skills and education.

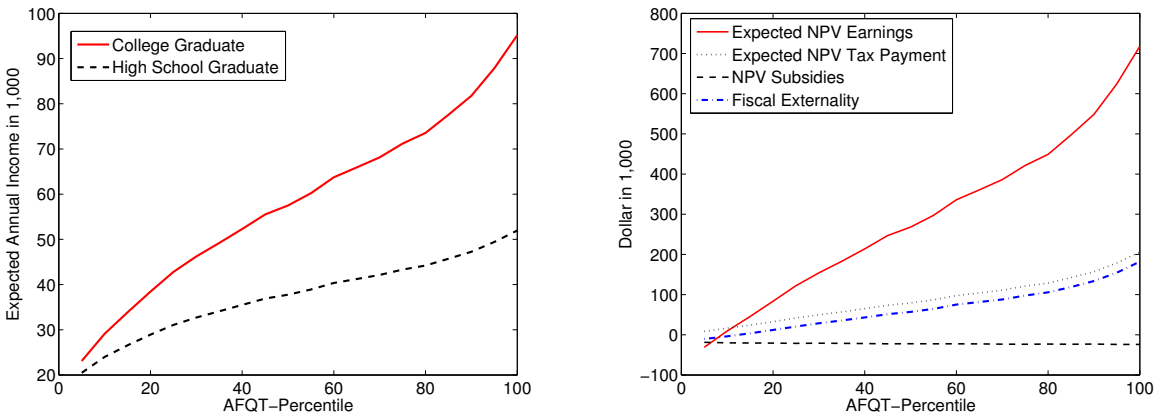
This procedure gives us the mean of log incomes as a function of an individual's AFQT-score and education level. Based on that, we then calculate the respective average annual income over the life cycle for each AFQT-score and education level. We assume errors are normally distributed, so income is distributed log-normally. To determine the second moment of this log-normal distribution across education and innate ability levels, we use the sample variances of the error terms from (6) for each education level.

Top incomes are underrepresented in the NLSY as in most survey data sets. Following common practice in the optimal tax literature (Piketty and Saez 2013), we therefore append Pareto tails to each income distribution, starting at incomes of \$350,000. We set the shape parameter a of the Pareto distribution to 2 for all income distributions.[26] Figure 1(a) shows the

[25] Note that this definition also includes high school dropouts and individuals with community college degrees. We also worked with different specifications but our main results were not significantly affected.

[26] Diamond and Saez (2011) find that starting from $\approx \$350,000$ the Pareto parameter is constant and 1.5. Since their data are for 2005 and our data are also for earlier periods, we choose a Pareto parameter of 2 because top incomes were less concentrated earlier. The rationale for having the Pareto parameter independent of education and innate ability is that we estimated this parameter separately for each education group, obtaining

expected annual before tax income as a function of the AFQT (in percentiles) for both education levels and clearly demonstrates the complementarity between innate ability and education, which has also been highlighted in previous papers (Carneiro and Heckman 2005). The red bold line in Figure 1(b) shows how this translates into an expected NPV difference in lifetime earnings. As was argued in the theoretical section, the returns to education play an important role for the fiscal effects of an increase in college enrollment. The additional tax payment (again in NPV) is clearly increasing in AFQT (black dotted line). To get the overall impact on the government budget, subsidies have to be subtracted, which are given by the black dashed-line. Subsidies are increasing in ability which reflects the fact individuals with higher ability currently obtain higher scholarships (merit-based financial aid), which we elaborate in Section 4.5. The net impact on public funds is given by the blue dashed-dotted line.



(a) Expected Annual Income

(b) NPV Income and Fiscal Externality

The last step consists of calibrating the respective skill/wage distribution from the income distributions by using the approximation of the US tax system from Guner, Kaygusuz, and Ventura (2013) and exploiting the first-order condition of individuals as pioneered by Saez (2001) and done in many papers after that.

4.4 Estimation of Parental Transfer Functions

In the NLSY97 we can observe the amount of transfers an individual obtained from its parents as well as family income. We take the constructed variable for parental transfers from Johnson (2013), who also takes into account the value of living at home as part of the parental transfer,

estimates very close to 3 in both cases. Additionally, we tested for any systematic patterns for a direct relationship between this parameter and θ without finding any robust or significant patterns. The next consideration is the well-known problem that high incomes are severely under-sampled in survey data as the NLSY or the PSID. We choose 2 as explained in the beginning of this footnote.

Table 2: Transfer Equation

	Parental Income	College	Dependent Children
Coefficient	.3136***	.5829***	-.0667**
Standard Error	(.0449)	(.0563)	(.0329)

N=3,238. Robust standard errors. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

for those individuals who cohabitate with their parents. We take yearly averages of those transfers for the ages 19-23. The sample average is $6,703. We estimate the following equation:[27]

$$log(tr_i) = \alpha^{tr} + \beta_1^{tr} log(\mathcal{I}_i) + \beta_2^{tr} co_i + \beta_3^{tr} depkids_i + \varepsilon_i^{tr},$$

where *depkids* is the number of dependent kids living in the household of the parents. The coefficients are provided in Table 2. A 1% increase in parental income increases parental transfers by 0.31% and college graduates receive transfers that are 79% ($exp(.5829) - 1$) higher than for high school graduates. Note that this implies that the absolute increase of parental transfers because of going to college is higher for high income kids. Johnson (2013) and Winter (2014) have argued that it is crucial to take this effect into account to explain the large impact of parental income on college enrollment and completion.

Besides transfers that individuals receive during that time, they can also have some assets when they decide to study. In the NLSY97, information is provided on individual net worth at age 20. Certainly, this is not the best number for our purposes since it is highly influenced by choices at ages 18 and 19. Since it gives our quantitative model a better fit concerning the importance of parental income, we nevertheless take this noisy measure into account. To see how net wealth varies with parental income, we estimate the regression:

$$w_i = \alpha^w + \beta^w \mathcal{I}_i + \varepsilon_i^w.$$

We find a gradient for parental income of .127 (0.02) and an intercept of $7,950 (1164).

4.5 Estimation of Grant Receipt

In practice, grants and tuition subsidies are provided by a variety of different institutions. Pell grants, for example, are provided by the federal government. In addition, there exist various state and university programs. To make progress, similar to Johnson (2013) and others, we go on to estimate the probability of grant receipt and the expected sum of the yearly grant directly from the data.

[27] We also estimated models with an interaction term between log parental income and college graduation. The coefficient on the interaction term is statistically insignificant though.

Table 3: Probit Estimation of Grants

	Parental Income	Black
Coefficient	-4.47e-06 ***	.2804***
Standard Error	(4.81e-07)	(.0623)

N=3,897. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

Table 4: OLS for Grants

	Parental Income	Black	AFQT
Coefficient	-.2405***	.3709 ***	2.1203 ***
Standard Error	(0.0586)	(0.1143)	(0.3493)

N=968. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

Since many students do not receive any grants at all, we first estimate a binary model of whether an individual receives a grant.

$$P(\mathcal{G}_i > 0) = Prob(Y_i^* > 0)$$

where

$$Y_i^* = \alpha^{gr1} + \beta_1^{gr1}\mathcal{I}_i + \beta_2^{gr1}black_i + \varepsilon_i^{gr1}$$

Interestingly, neither family size nor ability have a statistically significant effect at the extensive margin of grant receipt. As Table 3 reveals, blacks and low income children are more likely to obtain grants which captures affirmative action and need-based grants.

Next, we estimate the amount of grants conditional on receiving grants:

$$log(gr_i) = \alpha^{gr2} + \beta_1^{gr2}log(\mathcal{I}_i) + \beta_2^{gr2}black_i + \beta_3^{gr2}AFQT_i + \varepsilon_i^{gr2}.$$

Interestingly, family size is again not statistically significant. Besides grant generosity being need-based and in favor of blacks, generosity is also merit-based as $\hat{\beta}_3^{gr2} > 0$. The implied average grant is \$1,153 per year academic year.

4.6 Estimation of Psychic Costs

Based on the estimated reduced form relationships, we can calculate two value functions for each individual in the data: V_{co}^i and V_{hs}^i. These value functions are determined by innate ability and parental income.[28] In line with the empirical literature, we assume that the decision to

[28] As parental transfers also depend on the variable *depkids*, we set the variable *depkids* to its sample mean when calculating the implied transfers for an individual as a function of parental income. The same is true for the variable *black* because it determines grant receipt. Given the correlation of *black* with parental income and

Table 5: Probit Estimation of College Graduation

	AFQT	Father's Education	Mother's Education
Coefficient	.0140***	0.0832***	0.0504***
Standard Error	(.0011)	(.0097)	(.0105)

N=3,897. * p ≤ 0.10, ** p ≤ 0.05, *** p ≤ 0.01.

go to college is also influenced by heterogeneity in preferences for college. We assume that these psychic costs are determined by parental education and by innate ability – see Cunha, Heckman, and Navarro (2005), among others.[29] These assumptions give us the following binary choice model

$$P(co_i = 1) = Prob(Y_i^* > 0)$$

where

$$Y_i^* = V_{co}^i - V_{hs}^i + \beta_1^{pc} + \beta_2^{pc} * AFQT_i + \beta_3^{pc} S_i^{father} + \beta_4^{pc} S_i^{mother} + \varepsilon_i^{pc}$$

and where $\varepsilon_i^{pc} \sim N(0, \sigma)$. We restrict the coefficient on the difference in the value function to be one, as utility is our unit of measurement. For the power of the estimation, however, this is no restriction as a binary choice model always leaves one degree of freedom in parameter choice. As expected, all the variables have a positive and significant impact on the college choice.

Based on these estimations, we can calculate the estimated psychic costs for each individual:

$$\hat{\kappa}_i = \hat{\beta}_1^{pc} + \hat{\beta}_2^{pc} AFQT_i + \hat{\beta}_3^{pc} S_i^{father} + \hat{\beta}_4^{pc} S_i^{mother} + \hat{\varepsilon}_i^{pc}.$$

where $\hat{\varepsilon}_i^{pc} \sim N(0, \hat{\sigma})$. We draw 1,000 values for each ε_i and then fit a normal distribution of κ conditional on innate ability and parental income. Finally, we are equipped with the joint distribution of parental income, innate ability and psychic costs.

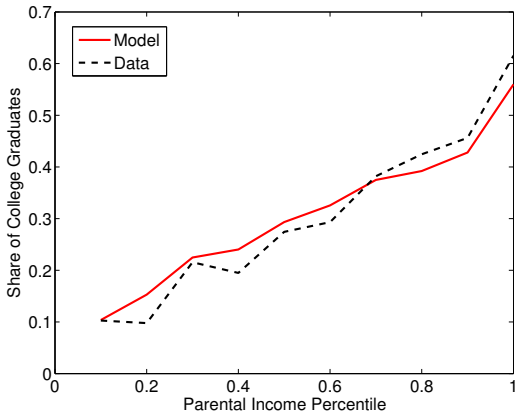
4.7 Model Performance

In order to assess the suitability of the model for policy analysis, we look at how well it replicates well known findings from the empirical literature and especially quasi-experimental studies. Note that the performance is also evaluated on moments which were not targeted in the estimation.

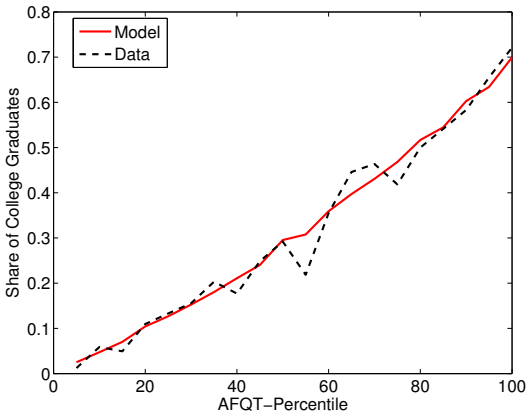
AFQT, we do not take the overall sample mean but rather the respective share of blacks for each combination of parental income and AFQT.

[29] The literature also suggests that individuals that grew up in urban areas are more likely to go college. The coefficient did not turn out as significant in our estimation and we therefore do not include it in our analysis. The inclusion of the variable does not affect any of our results.

Graduation Shares. Figure 1 illustrates graduation rates as a function of parental income and AFQT in percentiles respectively. The bold lines indicate results from the model and the dashed lines are from the data. In line with our result in the above paragraph, we slightly underestimate the parental income gradient. The correlation between AFQT and college graduation, however, is fitted well. The overall number of individuals with a bachelor degree is 30.56% in our sample and 30.95% in our model. Data from the United States Census Bureau are very similar: the share of individuals aged 25-29 in the year 2009 holding a bachelor degree is 30.6% – this comes very close to our data, where we look at cohorts born between 1980 and 1984.



(c) Graduation Rates and Parental Income



(d) Graduation Rates and AFQT

Figure 1: Removing Borrowing Constraints

Responsiveness of Graduation to Grant Increases. There is a large and established literature on the impact of increases in grants or decreases in tuition on college enrollment. Kane (2006) and Deming and Dynarski (2009) survey the literature. The estimated impacts of a $1,000 increase in yearly grants (or a respective reduction in tuition) on enrollment ranges from 1-6 percentage points, depending on the policy reform and research design. Numbers differ since some of the evaluated programs were targeted towards low income groups and others were not, and sometimes the higher amount of grants was associated with a lot of paperwork, which might create selection. The majority of studies arrive at numbers between 3 and 5 percentage points, however. As our model is a model of college graduation instead of college enrollment, the numbers are not directly comparable for two reasons: (i) not all of the newly enrolled students will indeed graduate with a bachelor's degree, (ii) some of the newly enrolled students enroll in community colleges and (iii) students that have enrolled also for lower grants are less likely to drop out of college. Relatively little is known about (iii). Concerning (i), we know that in the year 2000 roughly 66% of newly enrolled students enroll in 4-year institutions (Table 234

of Snyder and Dillow (2013)). Of those 66%, only slightly more than half should be expected to graduate with a bachelors degree. In Section 5.3, we estimate that the dropout probability of the marginal students in our model is 45%. However, of those initially enrolled at two year colleges, also 10% graduate with a bachelors degree (Figure 6 of Shapiro et al. (2012)).[30] Thus, translating the 3-5 percentage points increase in enrollment into numbers for graduation rates, we get 1.2-2 percentage points when taking into account (i) and (ii). Taking into account (iii) would yield slightly higher numbers, however, there is no strong empirical evidence about this effect that would guide us about the quantitative importance.

In order to test our model, we increase the grant for each individual by $1,000: bachelor degree completion rises by 1.48 percentage points. This number clearly lies within the reasonable range of estimates of 1.2-2.

A more recent study by Castleman and Long (2013) looks at the impact of grants targeted to low income children. Applying a regression-discontinuity design for need-based financial aid in Florida (Florida Student Access Grant), they find that a $1,000 increase in yearly grants for children with parental income around $30,000 increases enrollment by 2.5 percentage points. Interestingly, they find an even larger increase in the share of individuals that obtain a bachelor degree after 6 years by 3.5 percentage points. After 5 years the number is also quite high at 2.5 percentage points. These results show that grants can have substantial effects on student achievement conditional on enrollment.

Importance of Parental Income. It is a well known empirical fact that individuals with higher parental income are more likely to receive a college degree, see also Figure 1(c). However, it is not obvious whether this is primarily driven by parental income itself or variables correlated with parental income and college graduation. Using income tax data and a research design exploiting parental layoffs, Hilger (2014) finds that a $1,000 increase in parental income leads to an increase in college enrollment of 0.43 percentage points. Using a similar back of the envelope calculation as in the previous paragraph – i.e. that a 1 percentage point enrollment increase leads to a 0.40 percentage points increase in graduation rates – this implies an increase in graduation rates of .17 percentage points.

To test our model, we increased parental income for each individual by $1,000 and obtained increases in bachelors completion by 0.10 percentage points. In line with Hilger (2014), our model predicts a small effect of parental income, even a bit smaller than that of Hilger (2014). A likely explanation for why our model slightly understates the importance of parental income is that we assume that students cannot borrow on the private market. By assuming additional private borrowing that depends on parental income, we can get closer to the number of Hilger (2014). Given that there is no guidance from the empirical literature about how borrowing

[30] The numbers are for students initially enrolled in 2006 and therefore do not directly apply.

constraints might vary with parental income, we chose to stick to the case without private borrowing for our benchmark case.

The College Wage Premium and Marginal Returns. The college-earnings premium in our model is 100%, i.e. the average income of a college graduate is twice as high as the average income of a high-school graduate. As our earnings data are for the 1990s and the 2000s, this is well in line with empirical evidence in Oreopoulos and Petronijevic (2013); see also Lee, Lee, and Shin (2014). Doing the counterfactual experiment and asking how much the college graduates would earn if they had gone to college, we find that the returns to college are 64.40%. This implies a return of 12.43% for one year of schooling, which is in the upper half of the range of values found in Mincer equations (Card 1999, Oreopoulos and Petronijevic 2013)

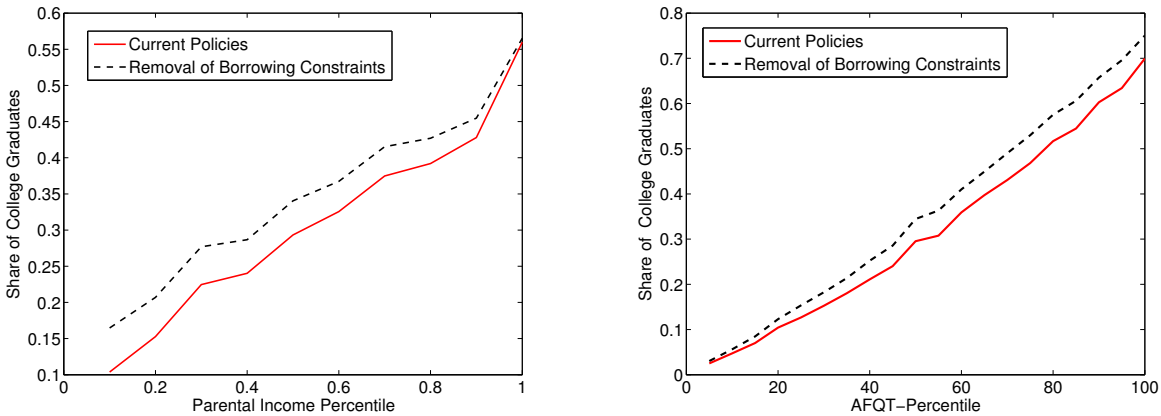
The more important number for our analysis is the return to college for *marginal students*. We find it to be slightly lower at 59.71%, which implies a return to one year of schooling of 11.71%. This reflects that marginal students are of lower ability on average than inframarginal students and also is in line with Oreopoulos and Petronijevic (2013). The probably cleanest study to infer returns for marginal students is Zimmerman (2014). In his study marginal refers to the *academically marginal* around a GPA admission cutoff. He finds returns of about 9.9% per year.[31] However, his number refers to the *academically marginal* students (implying a GPA of 3), whereas in our thought experiment we refer to those students who are marginal w.r.t. to a small change in financial aid – these students are likely to be of higher ability than the academically marginal students. We explore this issue and make use of the fact that the NLSY also provides GPA data. In fact our model gives a return to college of 52.29% for students with a GPA in the neighborhood of 3, which implies a Mincer return of 10.51% for one year of schooling – which comes very close to the 9.9% from Zimmerman (2014).

Finally, in these calculations we do not account for differing rates of unemployment and disability insurance rates. Both numbers are typically found to be only half as large for college graduates (See Oreopoulos and Petronijevic (2013) for unemployment and Laun and Wallenius (2013) for disability insurance). Further, the fiscal costs of Medicare are likely to be much lower for individuals with college degree. Lastly, we assume that all individuals work until 65 not taking into account that college graduates on average work longer (Laun and Wallenius 2013). These facts would strengthen our case for an increase in college subsidies in Section 5.

The Role of Borrowing Constraints. To assess the importance of borrowing constraints, we completely remove them to ask by how much graduation increases. The overall increase is

[31] He finds gains of 22% to obtain four-year college admission, which should be compared to the return of community colleges, which are the most frequent outside options for those students and take on average about 2 year less to complete. In addition, his findings are for earnings around 8 and 14 years after high school completion. Given that college students have a steeper earnings profile (see, e.g., Lee et al. 2014), these numbers are likely to underestimate the return to lifetime earnings.

neither small nor large: enrollment increases by 4.10 percentage points from 30.95% to 35.05%. This value is in the realm of values the literature has found, see, e.g., Johnson (2013) and Navarro (2011). As Figure 2(a) reveals, the removal of borrowing constraints has larger effects for low income children. Figure 2(b) illustrates the importance of borrowing constraints for individuals with different innate abilities. Naturally, individuals with high ability have the strongest need for more borrowing because of high expected future earnings.



(a) Borrowing Constraints and Parental Income

(b) Borrowing Constraints and AFQT

Figure 2: Removing Borrowing Constraints

5 College Policy Reforms

We have outlined the formal framework in Section 3 and described the parameterization and estimation in Section 4. In this section, we analytically *and* quantitatively assess the revenue consequences of various college subsidy reforms. In Section 5.1, we assess to what degree increasing tuition subsidies or college grants is self-financing through higher tax revenue in the future. In Section 5.2 we ask this question for need and merit-based grants. In Section 5.3, we discuss how our results would change if individuals can drop out of college. Finally, in Section 5.4 we show that not taking into account general equilibrium effects seems to be a minor problem for our results.

5.1 Uniform Increase of Education Grants

We first derive the effects on government revenue of marginally increasing all grants by $d\mathcal{G}$. First of all, a grant increase has a *mechanical* effect on revenue because additional money is paid to all students. This impact on the government budget is given by

$$M^{\mathcal{G}} = -d\mathcal{G} \times \beta^{co1} G^{co}(\overline{\omega}), \tag{7}$$

recall that $G^{co}(\overline{\omega})$ (as defined in (4)) is the share of college graduates. Everything is multiplied by β^{co1} because we define the grant increase such that it is equally increased in all years during college.

Apart from this mechanical revenue effect, the small grant increase also has a behavioral effect on public funds. Higher grants induce some students to go to college that would not have gone otherwise – we label these students as marginal students. These marginal students matter for the government budget since subsidies have to be paid to them and since they will pay a different amount of taxes over their life cycle. This change depends on innate ability and parental income because grants depend on ability as well as parental income and additional tax revenue depends on ability:

$$\Delta\mathcal{T}(\theta,\mathcal{I}) = \beta^{co2}\int_{\Omega}\mathcal{T}^{co}(y_{co}(\omega))dG^{co}_{\theta}(\omega) - \beta^{co1}\mathcal{G}(\theta,\mathcal{I}) - \beta^{hs}\int_{\Omega}\mathcal{T}^{hs}(y_{hs}(\omega))dG^{hs}_{\theta}(\omega).$$

The impact on the government budget of the increase in college graduation – that we call the college graduation effect – is given by:

$$CG^{\mathcal{G}} = d\mathcal{G} \times \int_{\Theta}\int_{\mathbb{R}_+}\Delta\mathcal{T}(\theta,\mathcal{I})\xi^{\mathcal{G}}(\theta,\mathcal{I})H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))dK_{\theta}(\mathcal{I})dF(\theta), \tag{8}$$

where $\xi^{\mathcal{G}}(\theta,\mathcal{I})$ is the semi-elasticity of college graduation with respect to grants as defined in Section 3.3.1. The overall impact on public funds of increasing $\mathcal{G}(\theta,\mathcal{I})$ is then given by:

$$\Gamma^{\mathcal{G}} = CG^{\mathcal{G}} + M^{\mathcal{G}}. \tag{9}$$

If $\Gamma^{\mathcal{G}}$ is negative, this implies that increasing public subsidies is not a free lunch. Otherwise, increasing subsides is (dynamically) self-financing as the additional tax revenue outweighs the direct fiscal costs of the increase in subsidies. Dividing by the negative of the mechanical effect, one gets

$$\gamma^{\mathcal{G}} = \frac{\Gamma^{\mathcal{G}}}{\beta^{co1}d\mathcal{G}G^{co}(\overline{\omega})}. \tag{10}$$

This expression can be interpreted as the rate of return on one dollar invested in additional college subsidies. If it takes the value .2, it says that the government gets \$1.20 in additional tax revenue for one marginal dollar invested into college subsidies. If it is -.5, it implies that the government gets 50 Cents back for each dollar invested. Inserting (8) and (9) into (10) gives

$$\gamma^{\mathcal{G}} = \frac{\int_{\theta}\int_{\mathcal{I}}\Delta e(\theta,\mathcal{I})\Delta\mathcal{T}(\theta,\mathcal{I})d\mathcal{I}d\theta}{\beta^{co1}G^{co}(\overline{\omega})} - 1, \tag{11}$$

where $\Delta e(\theta,\mathcal{I}) \equiv \xi^{\mathcal{G}}(\theta,\mathcal{I})H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))k_{\theta}(\mathcal{I})f(\theta)$ is the mass of marginal students of type $(\theta,\mathcal{I})$.

Next, define the normalized overall share of marginal students $\bar{\Delta}e = \int_\theta \int_{\mathcal{I}} \Delta e(\theta, \mathcal{I}) d\mathcal{I} d\theta / \beta^{co1}$: It is normalized such that it referes to a marginal increase of $1/\beta^{co1}$ of grants, i.e. a grant increase that costs the government exactly one dollar for each inframarginal student. Further, we define the average fiscal externality for marginal students $\bar{\Delta}\mathcal{T} = \frac{\int_\theta \int_{\mathcal{I}} \Delta e(\theta,\mathcal{I}) \Delta\mathcal{T}(\theta,\mathcal{I}) d\mathcal{I} d\theta}{\int_\theta \int_{\mathcal{I}} \Delta e(\theta,\mathcal{I}) d\mathcal{I} d\theta}$. This simplifies notation to get

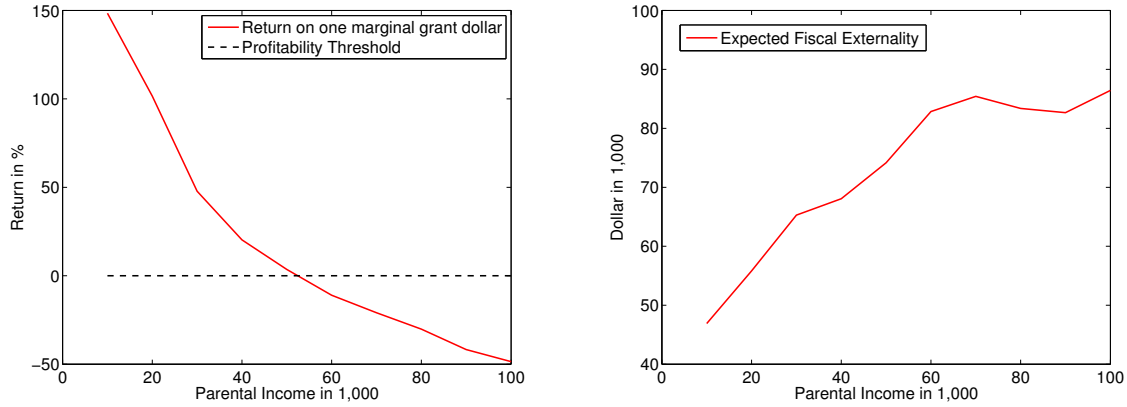
$$\gamma^{\mathcal{G}} = \frac{\bar{\Delta}e \times \bar{\Delta}\mathcal{T}}{G^{co}(\overline{\omega})} - 1. \tag{12}$$

Thus, whether the return on one dollar of additional college subsidies is positive depends on three measures: (i) the share of inframarginal students $G^{co}(\overline{\omega})$, (ii) the share of marginal students $\bar{\Delta}e$ and (iii) the average fiscal externality per marginal student $\bar{\Delta}\mathcal{T}$. This formula does not depend on the exact modeling specification but can instead be considered as a *sufficient-statistics* formula (Chetty 2009).

Having derived these effects, we now quantify (i)-(iii) using our quantified model which produces numbers for these three sufficient statistics that are very much in line with empirical evidence as argued in Section 4.7. The share of inframarginal students is 30.95%. The share of marginal students is .000381%. The fiscal externality per marginal student is \$70,940. This back of the envelope calculation reveals that $\gamma_R^{\mathcal{G}} = -.13$. Thus, the return on one marginal dollar invested into education grants is -13%. In other words, an increase in universal college subsidies is close to being self-financing; for every marginal invested dollar, the taxpayer gets \$0.87 back. Thus, increasing subsidies is to a large degree self-financing. It is straightforward to evaluate this formula for other values of these sufficient statistics. A lower responsiveness to a grant increase would make it smaller. It would increase for higher returns to education, a lower interest rate at which the government discounts and for a more progressive tax system. Finally, the lower the share of inframarginal students, the more profitable is a grant increase for the government. In the next subsection we elaborate the implications of a subsidy increase for different parental income levels and argue that the low share of inframarginal students for children with low parental income makes an increase in need-based financial aid particularly desirable.

5.2 Reforming Grants for Subgroups of Students

The previous subsection has revealed that three empirical measures together provide a sufficient statistic for the profitability of further public investment into college education. It is likely that these empirical measures vary a lot between different subgroups of potential college students. In the following paragraphs, we assess the profitability of increasing grants for certain subgroups: individuals with low parental income (need-based grants) and individuals with certain ability (merit-based) grants.



(a) Returns as a function of parental income

(b) Expected Fiscal Externality

Figure 3: Grant Increase for Different Parental Income Groups

5.2.1 Need-Based Grants

In our first experiment, we assess the profitability of public educational investment for students with different parental income. For example, if one thinks about increasing grants only for individuals with parental income below $\mathcal{I}^*$, formula (11) becomes

$$\gamma^{\mathcal{G}}_{\mathcal{I}^*} = \frac{\bar{\Delta} e_{\mathcal{I}^*} \times \bar{\Delta} \mathcal{T}_{\mathcal{I}^*}}{G^{co}_{\mathcal{I}^*}(\bar{\omega})} - 1, \tag{13}$$

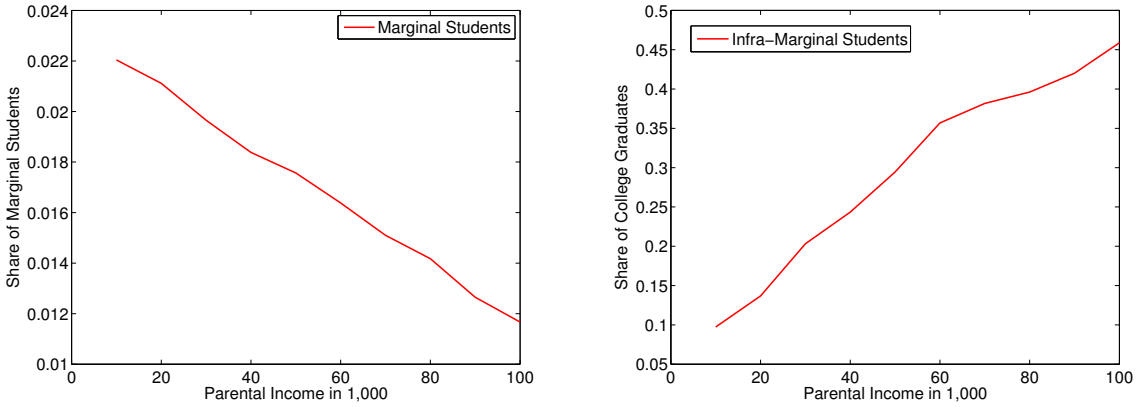
where all variables are adjusted to take into account that the reform only affects a subset of the population.[32] Crucially, the behavioral responses, marginal returns and student shares may differ a lot across parental income groups.

Our first experiment is to calculate this number for incomes up to \$50,000. We find $\gamma^{\mathcal{G}} = .37$ for this case. This is a striking result, suggesting very high returns to an expansion in targeted subsidies.

To better understand what the mechanisms are, we calculate γ for many points of the parent income distribution. The reforms here are "pointwise" at some parental income $\mathcal{I}$, meaning that we only increase grants for those students who have parental income $\mathcal{I}$. This transparently illustrates the mechanisms at work. Figure 3(a) illustrates these experiments. The profitability is very high for low incomes and strictly decreasing in parental income. Increasing grants for individuals with parental income around \$15,000 has a very high return of 150%, so that a \$1 increase in subsidies is estimated to increase tax revenue by \$2.50.

[32] Formally, the terms are defined s.t.

$$\gamma^{\mathcal{G}}_{\mathcal{I}^*} = \frac{\int_\theta \int_0^{\mathcal{I}^*} \Delta e(\theta, \mathcal{I}) \Delta \mathcal{T}(\theta, \mathcal{I}) d\mathcal{I} d\theta}{\int_\theta \int_0^{\mathcal{I}^*} \int_\Omega dG_\theta(\omega) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_\theta(\mathcal{I}) dF(\theta)} - 1.$$



(a) Marginal Students

(b) Inframarginal Students

Figure 4: Fiscal trade-off of increase in subsidy

Returns are positive up to an income of slightly above \$50,000 and become negative afterwards. This is a result which might go against a first intuition. One could have expected that the returns to college are larger for children with a better parental background, as those kids are, arguably, more apt to take advantage of the opportunities offered by a college education. Figure 3(b) shows that this intuition is indeed correct in our quantified model because the expected fiscal externality is clearly increasing in parental income stemming from the positive correlation between θ and $\mathcal{I}$. However, this effect is counteracted and dominated by the fact that the share of marginal over inframarginal students (i.e. the semi-elasticity of enrollment) is heavily decreasing in parental income. As can be seen in Figures 4(a) and 4(b), the share of marginal students (normalized to a "marginal" annual grant increase of \$1,000) mildly decrease by factor of roughly 2. Note that this in line Castleman and Long (2013) who find a strong responsiveness for children with low parental income. The share of inframarginal students varies more strongly and increases by a factor of roughly 4.5. Given that the fiscal costs of the reform are proportional to the share of inframarginal students, this implies that the costs of subsidizing low income kids are 4.5 times lower than the costs of subsidizing high income kids.

Our analysis shows that need-based education grants do not only favor social mobility and equity but are a good investment of society in future tax revenue. Remember that while the value of Δe for different income groups depends on our estimated model, the declining share of inframarginal students can be taken directly from the data. In fact, our results even *underestimate* the parental income gradient a bit, as discussed in Section 4.7. The fact that the result is to a large extent driven by those shares of inframarginal students across income groups should make the result particularly robust.

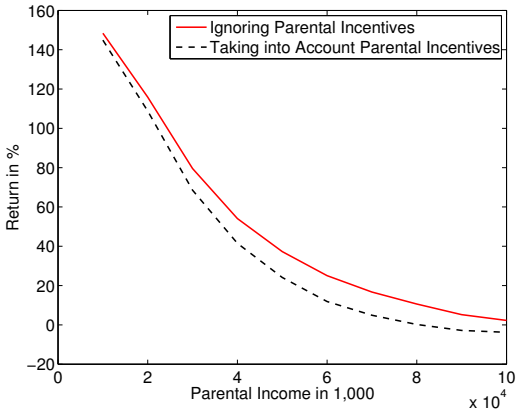


Figure 5: Fiscal trade-off of increase in subsidy

Policy Reforms and Parental Incentives. We now go back to the idea of increasing grants only for children with parental income below a certain threshold. Such reforms might impact parents' work incentives. If grants are increased for kids whose parental income is below a fixed threshold, then parents with incomes just above this threshold face an incentive to reduce their earnings. Making grants more need based, hence, increases effective marginal tax rates for parents. If parents work less, this will decrease tax revenue and therefore the profitability of such reforms will be reduced. But how important are those effects empirically? As we show in Appendix A.3, formula (13) needs to be adjusted as follows:

$$\gamma^{\mathcal{G}}_{\mathcal{I}^*} - \frac{h^{co}_{\mathcal{I}^*} \cdot \epsilon^{p}_{\mathcal{I}^*, 1-T'+\frac{\partial \mathcal{G}}{\partial \mathcal{I}}} \cdot \frac{T' - \frac{\partial \mathcal{G}}{\partial \mathcal{I}}}{1-T'+\frac{\partial \mathcal{G}}{\partial \mathcal{I}}} \cdot \mathcal{I}^*}{G^{co}_{\mathcal{I}^*}(\bar{\omega})},$$

where $h^{co}_{\mathcal{I}^*}$ is the mass of parents with income level $\mathcal{I}^*$ who have a child in college.

The additional term captures the reduced labor supply incentives of parents and the implied reduction in public funds. It is very similar to the budget effects of a marginal tax rate increase, triggered by a reduction in taxable in income (Piketty (1997), Saez (2001) and Golosov, Tsyvinski, and Werquin (2014)). $\epsilon^{p}_{\mathcal{I}^*, 1-T'+\frac{\partial \mathcal{G}}{\partial \mathcal{I}}}$ is the taxable income elasticity with respect to effective marginal tax rates for parents. The intuition here is that a small grant increase contingent on parental income implies an increase of the effective marginal tax rate for the affected parents. Although, to the best of our knowledge, there are no good estimates of how parental taxable income adjusts in response to need-based reforms, we can use estimates from the vast literature focusing on the taxable income elasticity. Finally, in the formula we have to use the total local labor wedge $T' - \frac{\partial \mathcal{G}}{\partial \mathcal{I}}$, consisting of the marginal income tax rate and the reduction in college grants.

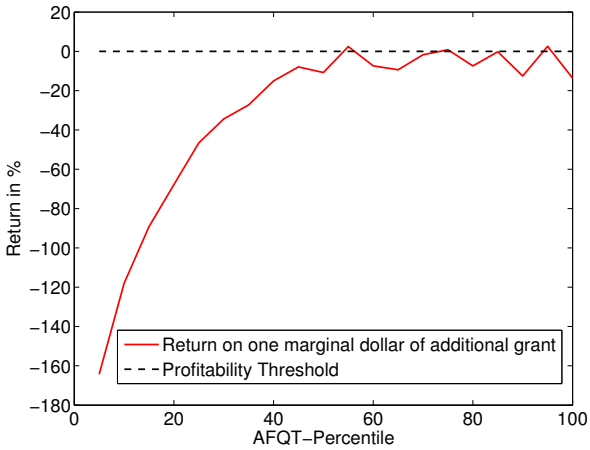


Figure 6: Returns as a function of AFQT

We set $\epsilon^{p}_{\mathcal{I}^*,1-T',\frac{\partial \mathcal{G}}{\partial \mathcal{I}}} = 0.33$. Since the relevant period where the parents could adjust their labor supply is only 4 periods, we choose a lower number than for the children, for whom we consider a lifetime labor supply decision.

Figure 5 illustrates the return on one dollar invested in grants for children with parental income below a certain threshold, taking into account parental incentives. The red bold line reflects the benchmark case (i.e. formula (13)) where parental incentives where not taken into account and the black dashed line reflects the case that they are taken into account. Returns are lowered, however, this effect is rather small. In particular, when the threshold is set at low parental income levels taking into account that parents might change their taxable income hardly affects the conclusions from the baseline case.

5.2.2 Merit-Based Grants

Another form of targeted student aid are merit-based grants. Most of merit-based scholarships are provided privately. Until 2011 the SMART grant program provided aid to needy students with a GPA of at least 3.0, thus it was partly merit-based. Figure 6 illustrates the return on one marginal dollar of aid as a function of the AFQT-score. As in Figure 3(a), we illustrate a hypothetical pointwise reform. Returns are positive if grant increases are targeted at individuals with AFQT-score above the 40th percentile. Afterwards it is almost flat. Two opposing effects cancel each other: whereas the fiscal externality is increasing in AFQT (see Figure 1(b)), the ratio of marginal over inframarginal students decreases.

5.3 Robustness Check I: Dropout

In this subsection, we elaborate how our results change if college dropout is taken into account as well. We base our calculations on empirical evidence on dropout behavior as well as own

Table 6: Dropout Probit

	$\mathcal{I}$	Father's Education	AFQT
Coefficient	-1.98e-06 ***	-.0622153 .***	-.0094686 ***
Standard Error	(6.47e-07)	(.0116132)	(.0013619)

N=1,921. * p $\leq$ 0.10, ** p $\leq$ 0.05, *** p $\leq$ 0.01.

estimates. Dropout changes the basic considerations from the previous subsections for two reasons: (i) subsidies are not only paid to those students who will graduate but also to those who drop out although for a shorter period of time. (ii) higher subsidies will not only induce individuals to study that will graduate but also some that will fail to graduate – whether this makes a subsidy increase fiscally more appealing depends on how large returns to 'some college' are.

We first look at effect (i) in isolation. Shapiro et al. (2012, Table 6) find that for the cohort which was first enrolled in a four year college in the fall of 2006, 62% graduated 6 years later. Thus, at most 38% never received a bachelor degree. Let's first assume that all marginal students indeed graduate. How does a first back of the envelope calculation change our result? The adjusted formula looks as follows

$$\gamma_D^{\mathcal{G}} = \frac{\Delta\bar{e} \times \Delta\bar{\mathcal{T}}}{G^{co}(\overline{\omega})\left(1 + \frac{1}{t_D}\frac{D_{infra}}{1-D_{infra}}\right)} - 1 = -.38, \tag{14}$$

where $\frac{1}{t_d}$ adjust for the fact that dropouts do not receive subsidies over 4.5 years, but less. We assume that dropouts on average spend two years in college, which is the average number dropouts spent in college in the NLSY97. This analysis reveals that taking into account dropout is not unimportant. We did not take into account fact (ii) in this first calculation, i.e. the impact of marginal students that drop out on the government budget. It is not obvious what value to take for the dropout probability of marginal students. In addition, marginal students also partly benefit from college. To tackle the first issue, we estimated the dropout probability in the NLSY97 as function of parental income, father's education and AFQT. The results are displayed in Table 6.

As expected, all these variables have a negative impact on dropping out of college.[33] Predicting the dropout probability of marginal and inframarginal students for our model yields a dropout probability of 36% for individuals with characteristics of our infa-marginal students and 45% for individuals that have characteristics of our marginal students. Thus the formula should now look like

[33]This is in line with reported statistics. Individuals with lower academic ability are more likely to dropout, see e.g. Chatterjee and Ionescu (2012). It has also been documented that parental characteristics play an important role, see e.g. Stinebrickner and Stinebrickner (2003).

$$\gamma_D^{\mathcal{G}} = \frac{\Delta\bar{e} \times \Delta\bar{\mathcal{T}}\left(\Delta\bar{\mathcal{T}} + \frac{D_{marg}}{1-D_{marg}}\Delta\bar{\mathcal{T}}^D\right)}{G^{co}(\overline{\omega})\left(1 + \frac{1}{t_D}\frac{D_{infra}}{1-D_{infra}}\right)} - 1 = -.30 \tag{15}$$

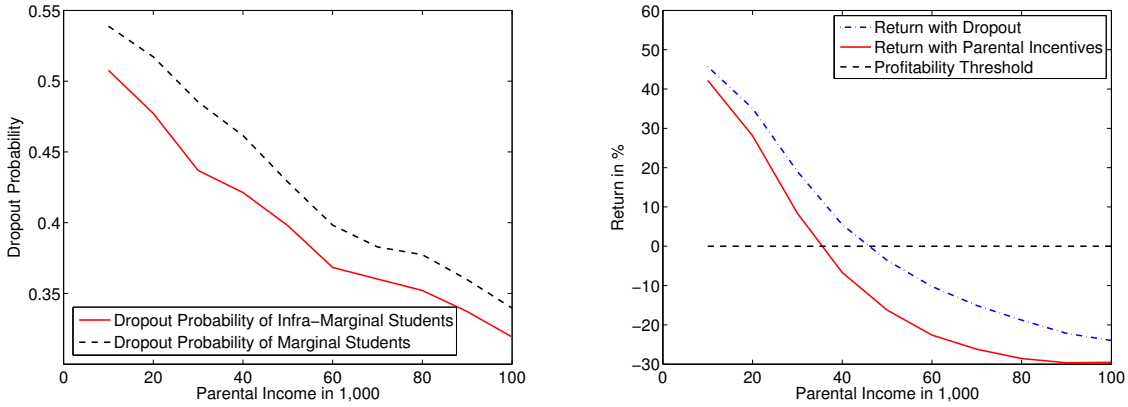
A question that remains is how to quantify $\Delta\bar{\mathcal{T}}^{dropout}$. According to Lee, Lee, and Shin (2014) the earnings premium for 'some college' was between 25% and 40% between 1980 and 2005. Based on that, we assume that the earnings increase from some college (i.e. 2 years in college before dropout) is 30% of that from graduating.[34] For this case, we obtain $\gamma^{\mathcal{G}} = -.30$ – thus even when taking dropout into account, the government gets back \$0.70 for each dollar invested.

Need-Based Grants We also investigate how our results for need-based grants change when taking into account college dropout. In particular, the results from Table 6 might raise doubt about the robustness of our result for the efficiency of need-based grants as low income children are more likely to drop out. We address this question and calculate the return on increasing subsidies for children with parental income below certain income levels again and take into account how the dropout probabilities of marginal and inframarginal students vary with income. Figure 7(a) shows how the dropout probability varies with parental income for marginal and inframarginal students of our model. The dropout probability clearly decreases with parental income. As Figure 7(b) (which is the dropout-equivalent to Figure 5) shows, this has indeed an effect on how the profitability of grant increases varies with parental income (See blue dash-dotted line). However, the effect is not too large and does not change much with respect to the main conclusion. We even look at a 'worst-case scenario' and take parental incentives into account as well. As the red bold line reveals even in this scenario increasing subsidies for children with parental income below \$35,000 is a free lunch; this corresponds to those children in the lowest parental income tercile.

5.4 Robustness Check II: General Equilibrium

Our analysis so far abstracted from general equilibrium effects. A rising share of college graduates is likely to decrease the returns to college. What does this imply for our findings? A first educated guess might be that it weakens the case for an increase in subsidies. If returns to college decline, wages decline not only for the marginal but also for inframarginal students. It turns out, however, that for the reforms we consider, general equilibrium effects seem to strengthen the argument for increasing subsidies rather than weakening it. The reason is that

[34] The numbers cited from Lee, Lee, and Shin (2014) do not take into account different ability of those who graduate and those who dropout. The number that we take (i.e. 30%) is conditional on ability. Given that those who dropout are typically of lower ability, we consider our choice of 30% as a lower bound.



(a) Dropout Probability

(b) Fiscal Return

Figure 7: Revenue Gains as a Function of Annual Income

the increase in college labor increases wages for high school graduates and therefore their contribution to public funds. There are approximately twice as many high-school than college graduates. As a consequence, the second effect (i.e. increases high-school wages) dominates and general equilibrium effects strengthen the case for an increase in subsidies.[35]

Following common practice, we assume a CES production function of the following form:

$$F(H, L) = ((A_l L)^\rho + (A_h H)^\rho)^{\frac{1}{\rho}}$$

where $\rho \in (-\infty, 1]$ and $\sigma = \frac{1}{1-\rho}$ is the elasticity of substation between college and non-college labor supply. For our model, we have

$$L = \int_\Theta \int_{\mathbb{R}_+} \int_\Omega y_{hs}(\omega) dG_\theta^{hs}(\omega) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_\theta(\mathcal{I}) dF(\theta)$$

and

$$H = \int_\Theta \int_{\mathbb{R}_+} \int_\Omega y_{co}(\omega) dG_\theta^{co}(\omega) \left(1 - H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I}))\right) dK_\theta(\mathcal{I}) dF(\theta).$$

It is simple to show that wages per efficiency unit of high school and college labor are given by

$$w_{hs} = A_l^\rho \left(A_l^\rho + A_h^\rho \left(\frac{H}{L} \right)^\rho \right)^{\frac{1-\rho}{\rho}}$$

and

$$w_{co} = A_h^\rho \left(A_h^\rho + A_l^\rho \left(\frac{H}{L} \right)^{-\rho} \right)^{\frac{1-\rho}{\rho}}.$$

[35] These considerations imply that we are not looking at how individual college decisions might change in the presence of general-equilibrium effects. This is an important long-run questions that we leave for future research. See Abbott, Gallipoli, Meghir, and Violante (2013) for a discussion on the long-term effects.

Without loss of generality, we set $w_{hs} = w_{co} = 1$ for our calibrated economy – it is w.l.o.g. since one cannot jointly identify the level of ω for workers and the wages for an efficiency unit of labor supply. We now repeat our analysis from above and ask what the revenue consequences are from marginally increasing grants for all (potential) college students. But now we assume that wages will change as a consequence of the increase in college labor supply. The consequence is that also tax payments of all individuals that do not change their college decision will change. The tax payments of college graduates will decrease as their wage decreases whereas the opposite is true for the high school graduates – for this simple exercise we assume that labor supply will be constant; as we argue below, endogenous labor supply would strengthen our results.

To compute how the number for $\gamma^{\mathcal{G}}$ changes, we therefore have to compute how the tax payments of those who do not change their behavior changes. Our adjusted formula thus reads as

$$\gamma_R^{\mathcal{G}} = \frac{\Delta\bar{e}\Delta\bar{\mathcal{T}} + G^{co}(\overline{\omega})\Delta\bar{\mathcal{T}}_{GE}^{co} + G^{hs}(\overline{\omega})\Delta\bar{\mathcal{T}}_{GE}^{hs}}{G^{co}(\overline{\omega})} - 1, \tag{16}$$

where $\Delta\bar{\mathcal{T}}_{GE}^{hs}$ and $\Delta\bar{\mathcal{T}}_{GE}^{co}$ reflect the average increase in the contribution to public funds due to general equilibrium effects on wages. We make the assumption that $\rho = .39$ which implies $\sigma = 1.64$ in accordance with Goldin and Katz (2009). We find that $\gamma^{\mathcal{G}} = -0.10$ for this case. Thus, general equilibrium effects increase this number by .03. Whereas tax payments from college graduates decrease, the opposite is true for high school graduates. As the share of high school graduates is larger than that of college graduates, it outweighs this effect even though high school wages increase less than college wages decrease and even though high school graduates face lower marginal tax rates given the progressivity of the tax code. Finally note that taking into account endogenous labor supply should strengthen our results: high school graduates would work more and college graduates would work less. Given that the wage increase effect of high school graduates dominated the wage decrease effect, it is unlikely that the opposite is true for labor supply responses. To conclude, taking into account that relative wages may change because of an increase of college educated labor supply does not affect the fiscal externality argument strongly and is likely to strengthen it. Another important question that we leave for future research is how decisions of future generations to go to college will be affected by the implied decline in the college wage premium and what this implies for policies.

6 Income Tax Reforms

In this section we derive formulas for the revenue effects of tax reforms and quantitatively assess them. We employ tax perturbation methods as in Piketty (1997), Saez (2001) and Golosov, Tsyvinski, and Werquin (2014). In Section 6.1, we consider reforms of the progressivity of the

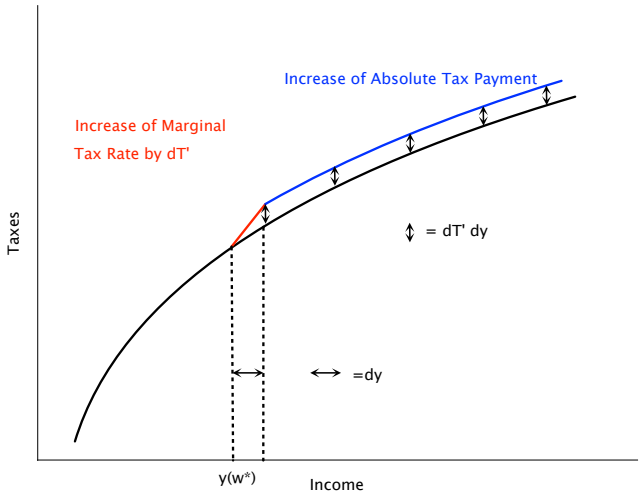


Figure 8: Tax Reform

tax system. Concretely, we ask how much revenue the government obtains from increasing marginal tax rates and how these numbers are affected by the college enrollment channel. We find that this channel is rather small as compared to the labor supply channel. In 6.2, we analyze the effects of education dependent tax reforms. For completeness, we also analyze the effects of lump-sum tax reforms in Appendix A.4.

6.1 Tax Progressivity, Excess Burden and College Education

We now consider the revenue effects of slightly changing marginal tax rates in small income intervals as originally considered by Piketty (1997) and Saez (2001) in a static framework and by Golosov, Tsyvinski, and Werquin (2014) in a dynamic framework. Figure 8 illustrates such a tax reform, where the marginal tax is increased by an infinitesimal amount $d\mathcal{T}'$ in an income interval of infinitesimal length $[y(\omega^*), y(\omega^*) + dy]$.

As a consequence of this reform, all individuals with $y > y(\omega^*)$ (and therefore $\omega > \omega^*$) face an increase of the absolute tax level of $d\mathcal{T}'dy$. The tax reform therefore induces a mechanical increase in tax revenue of

$$M(y(\omega^*)) = d\mathcal{T}'dy \left(\beta^{co2}\left(G^{co}(\overline{\omega}) - G^{co}(\omega^*)\right) + \beta^{hs}\left(G^{hs}(\overline{\omega}) - G^{hs}(\omega^*)\right)\right).$$

The increase in taxes for individuals with $\omega > \omega^*$ also changes incentives for college education. In fact, graduation will increase by:

$$d\mathcal{T}'dy \int_{\Theta}\int_{\mathbb{R}_+}\int_{\omega^*}^{\overline{\omega}} \xi^{ind}(\theta, \mathcal{I}, \omega) h_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_{\mathcal{I}}(\theta) dF(\theta).$$

where $\xi^{ind}(\theta, \mathcal{I}, \omega)$ is the semi-elasticity of college graduation as defined in Section 3.3.1. This increase in graduation has no first-order effect on welfare as these marginal individuals are just indifferent between obtaining a college degree or not. It has a first-order effect on the government budget which is given by:

$$CG(y(\omega^*)) = d\mathcal{T}'dy \int_{\Theta} \int_{\mathbb{R}_+} \Delta\mathcal{T}(\theta, \mathcal{I}) \int_{\omega^*}^{\overline{\omega}} \xi^{ind}(\theta, \mathcal{I}, \omega) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_{\mathcal{I}}(\theta) dF(\theta).$$

Note that discounting is already included in fiscal externality term $\Delta\mathcal{T}$, formally defined in Section 5.1. In addition, an increase in the marginal tax rate also affects labor supply behavior for individuals within the interval $[y(\omega^*), y(\omega^*) + dy]$. Individuals within this infinitesimal interval change their labor supply by

$$\frac{\partial y(\omega^*)}{\partial \mathcal{T}'} d\mathcal{T}' = -\varepsilon_{y,1-\mathcal{T}'} \frac{y}{1-\mathcal{T}'} d\mathcal{T}'. \tag{17}$$

Whereas this change in labor supply has no first-order effect on welfare on individual utilities by the envelope theorem, it has an effect on tax revenue. To obtain a measure for this tax revenue loss, define by $\tilde{g}^i(y) = g^i(\omega)\frac{\partial \omega}{\partial y}$ the density of income.[36] The mass of these individuals is then given by

$$\left(\tilde{g}^{co}(y(\omega^*)) + \tilde{g}^{hs}(y(\omega^*))\right) dy = \left(g^{co}(\omega^*) + g^{hs}(\omega^*)\right) \frac{\partial \omega}{\partial y} dy.$$

The overall impact on public funds (adjusted by period length and discounting) is therefore given by

$$LS(y(\omega^*)) = -\mathcal{T}' \left(g^{co}(\omega^*)\beta^{co2} + g^{hs}(\omega^*)\beta^{hs}\right) \frac{\partial \omega}{\partial y} dy \varepsilon_{y,1-\mathcal{T}'} \frac{y}{1-\mathcal{T}'} d\mathcal{T}',$$

which can be rewritten as

$$LS(y(\omega^*)) = -\mathcal{T}' \left(g^{co}(\omega^*)\beta^{co2} + g^{hs}(\omega^*)\beta^{hs}\right) \frac{\varepsilon_{y,1-\mathcal{T}'}}{1+\varepsilon_{y,1-\mathcal{T}'}} \frac{\omega^*}{1-\mathcal{T}'} d\mathcal{T}' dy$$

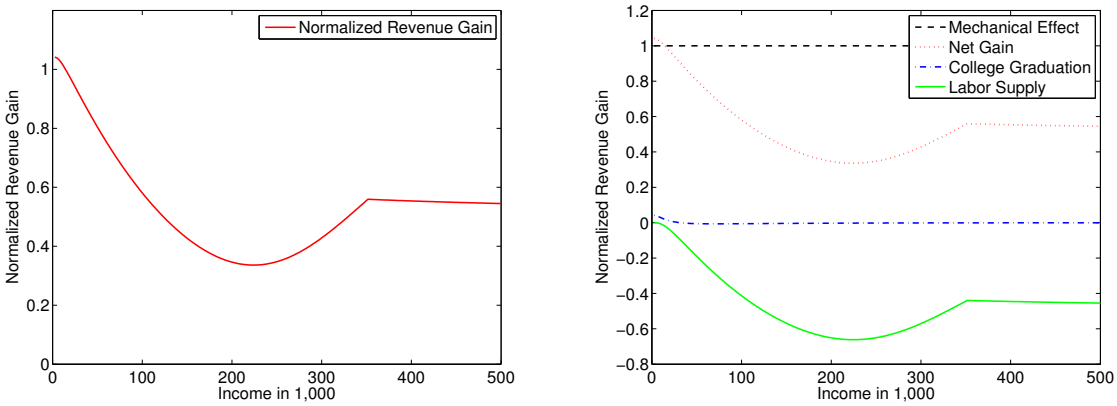
since $\varepsilon_{y,\omega} = 1 + \varepsilon_{y,1-\mathcal{T}'}$. The overall impact on welfare of the considered tax reform is thus given by

$$\Gamma(y(\omega^*)) = M(y(\omega^*)) + CG(y(\omega^*)) + LS(y(\omega^*)). \tag{18}$$

As in Section 5, we also define a normalized version of the revenue effect:

$$\gamma(y(\omega^*)) = \frac{\Gamma(y(\omega^*))}{d\mathcal{T}'dy \left(\beta^{co2}\left(G^{co}(\overline{\omega}) - G^{co}(\omega^*)\right) + \beta^{hs}\left(G^{hs}(\overline{\omega}) - G^{hs}(\omega^*)\right)\right)}. \tag{19}$$

[36] More formally, let $\tilde{G}^i(y)$ be the cdf of income. We then have $\tilde{G}^i(y(\omega)) = G^i(\omega)$. Taking the derivative with respect to y then yields $\tilde{g}^i(y) = g^i(\omega)\frac{\partial \omega}{\partial y}$.



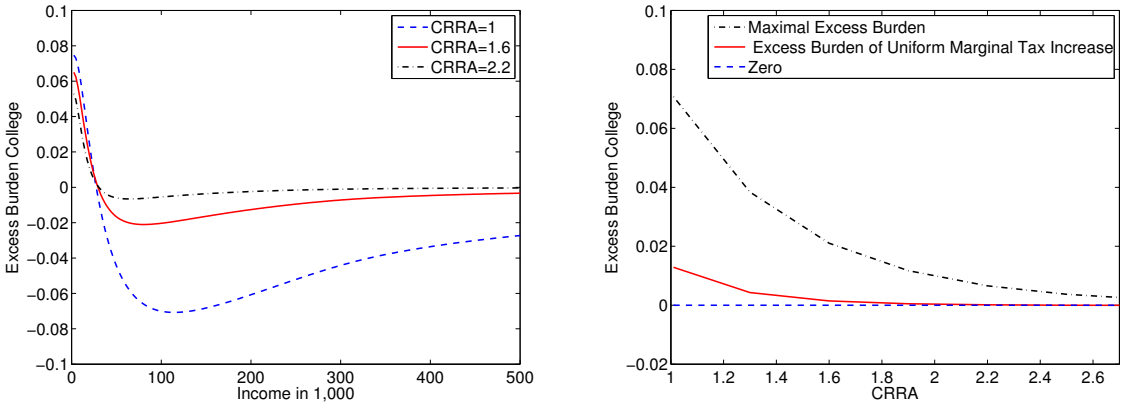
(a) Normalized Revenue Gain

(b) Decomposition

Figure 9: Revenue Gains as a Function of Annual Income

This term reads as follows: For each dollar that the government mechanically raises by increasing the marginal tax rate for individuals with income $y(\omega^*)$, it keeps $\gamma(\omega^*)$ dollars. The number therefore constitutes a quantification of Okun's leaky bucket (Okun 1975). The smaller this number, the leakier is the bucket – in other words: the more tax money is lost through behavioral responses. There is not enough empirical evidence to directly evaluate the college graduation term in (18) by using some numbers from the literature. Using our quantified version of the model, we obtain values for $\gamma(y(\omega^*))$ that are illustrated in Figure 9(a).

The intercept is slightly above one because the marginal tax rate at zero income has no labor supply effect. At the same time for marginal tax rates at very low income levels, the college graduation effect is positive by an income effect, see the discussion in Section 3.3.2. In general, the function is U-shaped. A similar quantitative result is obtained by Golosov, Tsyvinski, and Werquin (2014), who consider tax reforms in dynamic environments. As they elaborate, this is mainly driven by the way how hazard-rates of the income distribution vary with income. A related result is the U-shape of optimal marginal income taxes from the static literature (Diamond 1998, Saez 2001). Based on our results concerning the U-shape, we can carefully conclude that the endogeneity of college enrollment has no first-order effect on the shape of revenue gains. Figure 9(b) shows that the endogeneity also has no large effect on the size of the revenue gains. In this figure, the normalized revenue gain is decomposed into a normalized mechanical effect (which is by definition one), a normalized college graduation effect and a normalized labor supply effect. For income levels up to \$32,000 the college graduation effect is slightly positive. Through an income effect as described in Section 3.3.2, higher marginal tax rates increase incentives to acquire college education. For higher income levels it is negative although this can be hardly seen in the graph. At most the number is -.006, which implies that more than .6 Cents are not lost for each mechanically redistributed dollar – this is the



(a) Varying Risk Aversion

(b) Maximal and Average Values

Figure 10: Robustness of Excess Burden Result

case for increasing the marginal tax rate at income level \$71,502. For a uniform increase in marginal tax rates, the marginal excess burden is even only -0.02 Cents. We conclude that for our benchmark scenario the endogeneity of the college decision has a rather small impact on the efficiency costs of income taxes. As we show in Section 6.2, this changes if tax reforms are education dependent.

Robustness. Which parameters drive the above result that the additional excess burden is so low? Of those parameters which are preset, the coefficient of relative risk aversion and the borrowing maximum play an important role. We now illustrate what other value combinations for these parameters predict for the excess burden. In particular, we are interested in bounding the excess burden from above.

We fix the borrowing maximum at \$32,250, which is a case for which the excess burden was particularly strong. At this value, we typically find the strongest effects of tax reforms on graduation rates. We vary the CRRA between 1 and 2.7 – this is the range of values for which we obtain a responsiveness of the college margin to subsidies that is in line with empirical evidence.[37] For the values 1, 1.6 and 2.2, we illustrate the excess burden as a function of income in Figure 10(a). In all cases the college graduation effect is positive for incomes below $\approx$ \$30,000. Raising marginal tax rates below this income level encourages college education. For higher values, the numbers get negative – college education is discouraged if marginal tax rates for income above $\approx$ \$30,000 are increased. As can be seen this effect is decreasing in risk aversion. For $\gamma = 1$, up to 7 Cents for each mechanically raised Dollar are lost through the college channel; this number applies for marginal tax rate increases at \$114,610.

[37] Concretely, in these cases a \$1,000 increase in grants leads to an increase in graduation between 1 and 2%. For the discussion of the empirical evidence about this value see the discussion in the paragraph "Response of Graduation to Grant Increases" in Section 4.7.

The maximal value of the excess burden through the college education channel and the one of a uniform marginal tax rate increase are illustrated by the black dashed-dotted and the red bold line in Figure 10(b). The black dashed-dotted curve gives the maximal excess burden as a function of the CRRA coefficient. As just described, it is 7 Cents for $\gamma = 1$. It then quickly decreases and is below one if γ is above 2. Whereas this maximal value can get quite large, *on average* the excess burden through this channel is rather small. Average refers to a reform where the marginal tax rate is uniformly increased for all income levels. This is illustrated by the red bold line. Even for $\gamma = 1$, on average the additional loss for each mechanically raised tax dollar is below 1.2 Cents. We thus carefully conclude that on average the additional excess burden of income taxation caused by the college education channel is rather small.

6.2 Education Dependent Tax Reforms

For tax reforms, as illustrated in Figure 8, that only affect either college graduates or high school graduates, the mechanical effect is

$$\forall i = hs, co: \; M^i(y^i(\omega^*)) = d\mathcal{T}^{i'} dy \beta^i \left(G^i(\overline{\omega}) - G^i(\omega^*)\right).$$

The college graduation effect for education dependent marginal tax rate increases reads as

$$\forall i = hs, co: \; CG^i(y^i(\omega^*)) = d\mathcal{T}^{i'} dy \int_{\Theta} \int_{\mathbb{R}_+} \Delta\mathcal{T}(\theta, \mathcal{I}) \int_{\omega^*}^{\overline{\omega}} \xi^i(\theta, \mathcal{I}, \omega) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_{\mathcal{I}}(\theta) dF(\theta). \tag{20}$$

Since the semi-elasticities have unambiguous signs in this case, we know that this term is negative for $i = co$ and positive for $i = hs$ whenever the fiscal externality of going to college is positive for the average marginal student.

Finally, the education dependent labor supply effect is

$$\forall i = hs, co: \; LS^i(y^i(\omega^*)) = -\mathcal{T}^{i'} \beta^i g^i(\omega^*) \frac{\varepsilon_{y^i, 1-\mathcal{T}^{i'}}}{1 + \varepsilon_{y^i, 1-\mathcal{T}^{i'}}} \frac{\omega^*}{1 - \mathcal{T}^{i'}} d\mathcal{T}^{i'} dy,$$

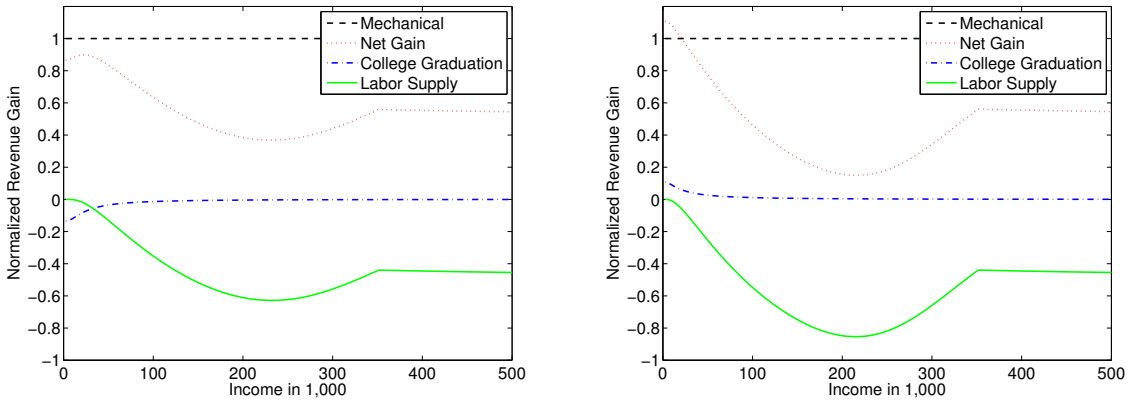
which follows from similar analytical steps as in the education independent case. These considerations yield the following overall effect on tax revenue

$$\forall i = hs, co: \; \Gamma^i(y_i(\omega^*)) = M^i(y^i(\omega^*)) + CG^i(y^i(\omega^*)) + LS^i(y^i(\omega^*)). \tag{21}$$

Again define the normalized version as

$$\gamma^i(y_i(\omega^*)) = \frac{\Gamma^i(y_i(\omega^*))}{d\mathcal{T}^{i'} dy \beta^i \left(G^i(\overline{\omega}) - G^i(\omega^*)\right)}.$$

These two functions and their decompositions are illustrated in Figure 11. The college grad-



(a) Tax Increases for College Graduates

(b) Tax Increases for High School Graduates

Figure 11: Revenue Gains as a Function of Annual Income

uation effect has a unique sign and is significantly larger than in the education-independent case for income levels up to \$50,000. When increasing taxes for college graduates, for each redistributed dollar up to 13 Cents are lost through the college margin in our benchmark case. For high school taxes, up to an additional 11 Cents for each mechanically raised dollar are gained via positive incentives for college education. These numbers can get much larger for other specifications. For the case of $\gamma = 1$ and a borrowing maximum of \$32,250, up 43 Cents of each mechanical raised Dollar (via taxes for college graduates) of tax revenue are lost via dampened incentives for college education. When increasing taxes for high school graduates solely, up to 26 Cents of additional tax revenue are obtained for each mechanically raised Dollar via the positive incentives for college education.

7 Optimal Policies

In this section, we analytically and quantitatively elaborate optimal policies with various degrees of sophistication. Our main question is by how much quantitative results from the theory of optimal nonlinear income taxation (Saez 2001) are altered by the college margin and how this depends on the level of subsidies. We therefore restrict subsidies to be independent of parental background and ability in this section. In Section 7.2, we derive an intuitive formula for the optimal college subsidy that holds in the presence of optimal *and* suboptimal taxes. We also simulate the optimal subsidy given the current US tax system. In Section 7.3, we derive an expression for the optimal marginal income tax rate that can well be related to the literature on optimal nonlinear taxation going back to Mirrlees (1971) and quantitatively assess it taking the current US education policies as given. In Section 7.4, we consider the case of setting income taxes and college subsidies jointly before we elaborate a more sophisticated system of education dependent taxes in Section 7.5.

For all quantitative explorations of optimal policies, we assume that the government can also overcome borrowing constraints. Alternatively one could have studied optimal policies in the presence of borrowing constraints. For this case, optimal policies would have played the additional role of relaxing borrowing constraints. Exploring this case as well would be a relatively straightforward extension. From a theoretical point of view, our formulas for optimal policies presented below also hold for binding borrowing constraints.

Robustness checks for our results in Sections 7.3 and 7.4 are provided in Appendix A.6, where different values of the CRRA coefficient are explored. It is shown that the most important results do not depend to heavily on that parameter value.

7.1 Social Objective and Constraints

We consider a social planner that designs $\mathcal{T}^{hs}(\cdot)$, $\mathcal{T}^{co}(\cdot)$ and $\mathcal{G}$. The planner's objective is assumed to be given by:

$$\int_{\Theta}\int_{\mathbb{R}_+}\int_{\underline{\kappa}}^{\tilde{\kappa}(\theta,\mathcal{I})} \left(V_{co}(\theta,\mathcal{I})-\kappa\right) d\tilde{H}_{\theta,\mathcal{I}}(\kappa)d\tilde{K}_{\theta}(\mathcal{I})d\tilde{F}(\theta)$$
$$+\int_{\Theta}\int_{\mathbb{R}_+}\int_{\tilde{\kappa}(\theta,\mathcal{I})}^{\overline{\kappa}} V_{hs}(\theta,\mathcal{I})d\tilde{H}_{\theta,\mathcal{I}}(\kappa)d\tilde{K}_{\theta}(\mathcal{I})d\tilde{F}(\theta), \tag{22}$$

where $\tilde{\kappa}(\theta,\mathcal{I})$ is defined by (3) and $V_{co}(\theta,\mathcal{I})$ and $V_{hs}(\theta,\mathcal{I})$ are defined by (2) and (1). $\tilde{H}_{\theta,\mathcal{I}}(\kappa)$, $\tilde{K}_{\theta}(\mathcal{I})$ and $\tilde{F}(\theta)$ denote the cumulated Pareto weights. If, e.g., $\tilde{H}_{\theta,\mathcal{I}}(\kappa)=H_{\theta,\mathcal{I}}(\kappa)$, $\tilde{K}_{\theta}(\chi)=K_{\theta}(\chi)\ \forall\ \theta,\chi$ and $\tilde{F}(\theta)=F(\theta)\ \forall\ \theta$, the social planner is Utilitarian.

Policies have to satisfy a government budget constraint:

$$\int_{\Theta}\int_{\mathbb{R}_+}\left(-\beta^{co1}\mathcal{G}+\beta^{co2}\int_{\Omega}\mathcal{T}^{co}(y_{co}(\omega))dG_{\theta}^{co}(\omega)\right)H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))dK_{\theta}(\mathcal{I})dF(\theta)$$
$$+\int_{\Theta}\int_{\mathbb{R}_+}\beta^{hs}\int_{\Omega}\mathcal{T}^{hs}(y_{hs}(\omega))dG_{\theta}^{hs}(\omega)\left(1-H_{\theta,\mathcal{I}}(\tilde{\kappa})\right)dK_{\theta}(\mathcal{I})dF(\theta)\geq 0. \tag{23}$$

Further, we denote as ρ the multiplier on (23), which reflects the marginal value of public funds.

We make use of welfare weights that are endogenous with respect to policies. We define the welfare weight of an individual of type $(\theta,\mathcal{I},\kappa,\omega)$ with $\kappa<\tilde{\kappa}(\theta,\mathcal{I})$ as[38]

$$\mathcal{W}(\theta,\kappa,\mathcal{I},\omega)=\frac{\beta^{co}}{\rho}U'\left(c_{co}(\theta,\mathcal{I},\omega)-v\left(\frac{y_{co}(\omega)}{\omega}\right)\right)\frac{\tilde{h}_{\theta,\mathcal{I}}(\kappa)\tilde{k}_{\theta}(\mathcal{I})\tilde{f}(\theta)}{h_{\theta,\mathcal{I}}(\kappa)k_{\theta}(\mathcal{I})f(\theta)}.$$

[38]For $\kappa>\tilde{\kappa}(\theta,\mathcal{I})$, we only would have to replace subscript co by hs.

Such a weight is endogenous because it depends on the marginal utility of consumption and therefore on policies. It can be read as follows: marginally increasing consumption for individuals of type $(\theta, \mathcal{I}, \kappa, \omega)$ increases welfare in units of public funds by $\mathcal{W}(\theta, \mathcal{I}, \kappa, \omega) \times g_\theta^{co}(\omega) h_{\theta,\mathcal{I}}(\kappa) k_\theta(\mathcal{I}) f(\theta)$. Or, equivalently, increasing consumption of $(\theta, \mathcal{I}, \kappa, \omega)$ by one dollar is as desirable as obtaining $\mathcal{W}(\theta, \chi, \omega) \times g_\theta^{co}(\omega) h_{\theta,\mathcal{I}}(\kappa) k_\theta(\mathcal{I}) f(\theta)$ additional dollars of public funds. We now also define two aggregated welfare weights. The welfare weights of all individuals with wage ω and with/without a college degree are given by:

$$\mathcal{W}^{co}(\omega) = \frac{1}{\rho} \frac{\int_\Theta \int_{\mathbb{R}_+} \int_{\underline{\kappa}}^{\tilde{\kappa}(\theta,\mathcal{I})} U'\left(c_{co}(\theta, \mathcal{I}, \omega) - v\left(\frac{y_{co}(\omega)}{\omega}\right)\right) g_\theta^{co}(\omega) d\tilde{H}_{\theta,\mathcal{I}}(\kappa) d\tilde{K}_\theta(\mathcal{I}) d\tilde{F}(\theta)}{\int_\Theta \int_{\mathbb{R}_+} \int_{\underline{\kappa}}^{\tilde{\kappa}(\theta,\mathcal{I})} g_\theta^{co}(\omega) dH_{\theta,\mathcal{I}}(\kappa) dK_\theta(\mathcal{I}) dF(\theta)} \tag{24}$$

and

$$\mathcal{W}^{hs}(\omega) = \frac{1}{\rho} \frac{\int_\Theta \int_{\mathbb{R}_+} \int_{\tilde{\kappa}(\theta,\mathcal{I})}^{\overline{\kappa}} U'\left(c_{hs}(\mathcal{I}, \omega) - v\left(\frac{y_{hs}(\omega)}{\omega}\right)\right) g_\theta^{hs}(\omega) d\tilde{H}_{\theta,\mathcal{I}}(\kappa) d\tilde{K}_\theta(\mathcal{I}) d\tilde{F}(\theta)}{\int_\Theta \int_{\mathbb{R}_+} \int_{\tilde{\kappa}(\theta,\mathcal{I})}^{\overline{\kappa}} g_\theta^{hs}(\omega) dH_{\theta,\mathcal{I}}(\kappa) dK_\theta(\mathcal{I}) dF(\theta)} \tag{25}$$

Based on that, the social marginal welfare weight of all individuals with wage ω is given by $\mathcal{W}(\omega) = \mathcal{W}^{co}(\omega) + \mathcal{W}^{hs}(\omega)$.

The above weights were defined for the working population. We also define the welfare weight of a college student of type $(\theta, \mathcal{I}, \kappa)$ by

$$\mathcal{W}^{st}(\theta, \mathcal{I}, \kappa) = \frac{1}{\rho} U'(c_e(\theta, \mathcal{I})) \frac{\tilde{h}_{\theta,\mathcal{I}}(\kappa) \tilde{k}_\theta(\mathcal{I}) \tilde{f}(\theta)}{h_{\theta,\mathcal{I}}(\kappa) k_\theta(\mathcal{I}) f(\theta)}.$$

Similar as the aggregated worker weights, we also define an aggregated student weight

$$\mathcal{W}^{st} = \frac{1}{\rho} \frac{\int_\Theta \int_{\mathbb{R}_+} \int_{\underline{\kappa}}^{\tilde{\kappa}(\theta,\mathcal{I})} U'(c_e(\theta, \mathcal{I})) d\tilde{H}_{\theta,\mathcal{I}}(\kappa) d\tilde{K}_\theta(\mathcal{I}) d\tilde{F}(\theta)}{\int_\Theta \int_{\mathbb{R}_+} \int_{\underline{\kappa}}^{\tilde{\kappa}(\theta,\mathcal{I})} dH_{\theta,\mathcal{I}}(\kappa) dK_\theta(\mathcal{I}) dF(\theta)}. \tag{26}$$

7.2 Optimal Education Subsidies

We start with a simple formula for the optimal college grant:

Proposition 1. *The optimal college grant $\mathcal{G}$ satisfies:*

$$\mathcal{G} = \frac{\int_\theta \int_{\mathbb{R}_+} \Delta\mathcal{T}(\theta, \mathcal{I}) \eta^{co}(\theta, \mathcal{I}) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_\theta(\mathcal{I}) dF(\theta)}{\beta^{co1} G^{co}(\overline{w}) \left(1 - \mathcal{W}^{st}\right).}, \tag{27}$$

where $\eta^{co}(\theta, \mathcal{I}) = \xi^{\mathcal{G}}(\theta, \mathcal{I}) \cdot \mathcal{G}$ is the elasticity of college graduation w.r.t. the education subsidy for individuals of type $(\theta, \mathcal{I})$ and $\mathcal{W}^{st}$ is the average social welfare weight of a student as defined in (26).

Proof. See Appendix A.5.1 □

The intuition is very similar to the intuition behind the revenue gains from increasing college subsidies outlined in Section 5.1: The numerator captures the *fiscal externality*, i.e. the gains from subsidizing college. The stronger individuals respond to subsidies (captured by the elasticity $\eta^{co}(\theta, \mathcal{I})$), the larger subsidies should be. In addition, the higher taxes are and the higher the returns to college are, the larger is $\Delta\mathcal{T}(\theta, \mathcal{I})$ and the larger subsidies should be.

The denominator reflects the cost of increasing college subsidies. *Ceteris paribus*, higher college graduation rates $G^{co}(\overline{w})$ are associated with a lower subsidy, as all graduates receive the subsidy. But this cost term is corrected by the factor $1 - \mathcal{W}^{st}$ since the planner also values the consumption of each marginal student by $\mathcal{W}^{st}$ in terms of public funds. Note also that this formula for the optimal education subsidy does not only apply for the optimal income tax but also for any given suboptimal income tax function.

The Role of Borrowing Constraints & Uncertainty. Formula (27) does not *explicitly* depend on the degree of uncertainty and the presence and tightness of borrowing constraints. Both aspects nevertheless influence the size of the optimal college grant $\mathcal{G}$. In this paragraph we argue that an increase in uncertainty and an increase in the tightness of borrowing constraints both are likely to increase the optimal level of college education subsidies. For the former, we think of a small mean-preserving spread of $G^{co}_{\theta}(\omega)$. For the latter, we think of a small increase in $\bar{s}$. Both changes in parameters will decreases the share of students. The former renders college less attractive because there is no full insurance and individuals are risk averse. The latter will trivially render college less attractive if borrowing constraints are binding. The decrease in the share of students in both cases decreases the marginal costs of the subsidy given in the numerator of (27); at least for the case that the social marginal welfare weight of college students is not too high, i.e. $\mathcal{W}^{st} < 1$.

An increase in uncertainty and a tightening of borrowing constraints also affects the marginal gain from increasing the subsidy by the behavioral response of enrollment w.r.t. grants. To better understand the impact, substitute for $\eta^{co}(\theta, \mathcal{I})$ in the numerator of (27), which yields

$$\int_{\theta}\int_{\mathbb{R}_+} \Delta\mathcal{T}(\theta, \mathcal{I})\mathcal{G}\frac{\partial\tilde{\kappa}(\theta, \mathcal{I})}{\partial\mathcal{G}} h_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I}))dK_{\theta}(\mathcal{I})dF(\theta) \tag{28}$$

where $\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{G}} = \beta^{co1}U'(c_e)$. Tightening binding borrowing constraints implies a higher value for $\frac{\partial\tilde{\kappa}(\theta,\mathcal{I})}{\partial\mathcal{G}}$ since c_e decreases. A mean-preserving spread will have no effect on c_e if individuals are borrowing constrained. If individuals are not borrowing constrained, it will imply a decrease of c_e – higher risk leads to an increase in precautionary savings. For arbitrary conditional distributions of psychic costs, it is in general unclear whether (28) increases or decreases. If these distributions are uniform, however, the above term unambiguously increases if borrowing

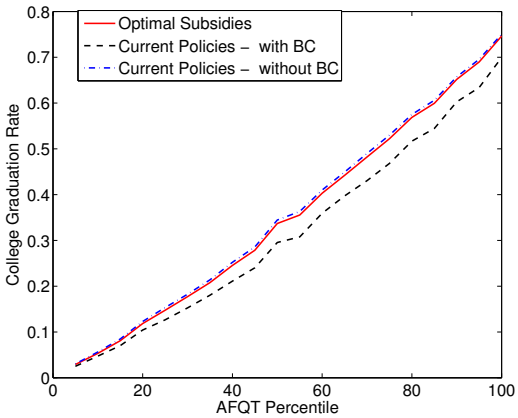


Figure 12: Optimal Education Subsidies for Given Income Taxes

constrained are tightened given that borrowing constraints are binding. It also increases if wage risk for college graduates is increasing given that borrowing constraints are not binding. To sum up, marginal gains from increasing the college grant are weakly increasing if borrowing constraints are tightened or college wage risk is increased. Since marginal costs of increasing subsidies are decreasing in both cases, the optimal college grant is therefore increasing in both cases.

Quantitative Exercise We look at the following scenario: the government cannot change marginal tax rates, however, it can adjust the lump sum element of the tax function and the college subsidy. Further, we assume that individuals can freely borrow. In this case the social planner slightly increases the lump-sum element to \$2,005 from \$1,938, which is the value of the transfer that the government can provide for current policies, but in the absence of borrowing constraints. Grants decrease from \$5,231 to \$3,905. The implied small decrease in enrollment is shown in Figure 12. The optimal Utilitarian graduation (holding marginal tax rates fixed) is 34.44%, and therefore slightly lower than the graduation rate given current policies absent borrowing constraints (35.05%). These results might seem at odds with our results in favor of higher subsidies in Section 5. Why is the optimal subsidy here lower than the current level of subsidies? The reason is that we are considering a case where individuals can freely borrow. The additional effects of larger subsidies on enrollment are therefore significantly lower. Increasing subsidies would therefore require to lower the lump sum transfer that individuals receive (i.e. increase $T(0)$). A Utilitarian planner does the opposite at the margin since the social marginal utility of college students (given that they can freely borrow) is relatively high. To provide more consumption for individuals with low income realizations (hence, insurance), the planner slightly increases the lump-sum transfer that individuals receive at the cost of reducing the

subsidy even though the reduction in subsidies comes with additional costs from the reduction in college graduation.

7.3 Optimal Income Taxes

We start by showing a formula for the optimal marginal income tax rate in the spirit of Diamond (1998) and Saez (2001):

Proposition 2. *Optimal education independent marginal tax rates satisfy*

$$\frac{\mathcal{T}'(y(\omega^*))}{1-\mathcal{T}'(y(\omega^*))} = \left(1+\frac{1}{\varepsilon(\omega^*)}\right)\times\left(\frac{\int_{\omega^*}^{\overline{\omega}}\left(\beta^{co2}\left(1-\mathcal{W}^{co}(\omega)\right)+\beta^{hs}\left(1-\mathcal{W}^{hs}(\omega)\right)\right)dG(\omega)}{\left(g^{co}(\omega^*)\beta^{co2}+g^{hs}(\omega^*)\beta^{hs}\right)\omega^*}\right. \tag{29}$$
$$\left.+\frac{\int_{\omega^*}^{\overline{\omega}}\int_{\Theta}\int_{\mathbb{R}_+}\Delta\mathcal{T}(\theta,\mathcal{I})\xi^{ind}(\theta,\mathcal{I},\omega)H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))dK_{\mathcal{I}}(\theta)dF(\theta)d\omega}{\left(g^{co}(\omega^*)\beta^{co2}+g^{hs}(\omega^*)\beta^{hs}\right)\omega^*}\right)$$

where $\mathcal{W}^{co}(\omega)$ and $\mathcal{W}^{hs}(\omega)$ are as defined in (24) and (25). $\varepsilon(\omega)$ is the labor supply elasticity as implicitly defined in (17) and the college semi-elasticities $\xi^{ind}(\theta,\mathcal{I},\omega)$ is as derived in Section 3.3.1.

Proof. See Appendix A.5.2 □

First, note that this formula holds for optimal as well as for suboptimal college subsidies. This formula for the optimal marginal tax rate differs from the formula of Diamond (1998) in two respects. First of all, it is adjusted for period length and discounting. Second, the term

$$\int_{\omega^*}^{\overline{\omega}}\int_{\Theta}\int_{\mathbb{R}_+}\Delta\mathcal{T}(\theta,\mathcal{I})\xi^{ind}(\theta,\mathcal{I},\omega)H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))dK_{\mathcal{I}}(\theta)dF(\theta)d\omega$$

shows up in the numerator. The formula is therefore related to the formulas of Saez (2002) and Jacquet, Lehmann, and Van der Linden (2013), where the extensive margin is due to labor market participation, or Lehmann, Simula, and Trannoy (2014) where the extensive margin captures migration. In these papers, the extensive margin is an unambiguous force towards lower marginal tax rates whenever workers pay more taxes than non-workers (or individuals that are on the margin of emigrating pay positive taxes). In contrast, the endogeneity of college enrollment does not necessarily lead to lower marginal tax rates as the additional term is ambiguous in its sign. First, we do not know the sign of $\Delta\mathcal{T}(\theta)$ in general. Second, we do not know whether higher taxes for individuals with $\omega > \omega^*$ indeed lead to lower college enrollment because of possibly counteracting income and substitution effects as outlined in Section 3.3.2. Whether and to what extent the endogeneity of college enrollment leads to lower optimal marginal tax rates is a quantitative question, which we address below.

If the upper tail of the earning distribution is Pareto distributed, one can derive a simple formula for the optimal top tax rate. Consistent with the empirical distributions we use, we assume that for both education groups above a high-income threshold y^T, earnings are Pareto distributed with common parameter a.

Proposition 3. *Assume that welfare weights for top incomes converge to a constant $\overline{\mathcal{W}}$. Then the optimal top marginal tax rate above income level y^T satisfies:*

$$\frac{\tau}{1-\tau} = \frac{1-\overline{\mathcal{W}}+\overline{CG}\cdot a}{1-\overline{\mathcal{W}}+\epsilon\cdot a},$$

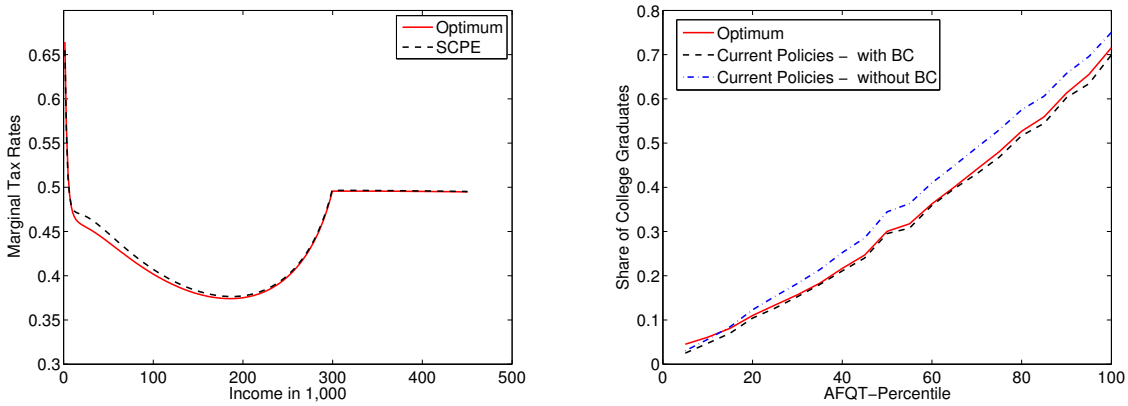
where $\overline{CG} = \frac{\int_\Theta \int_{\mathbb{R}_+} \Delta\mathcal{T}(\theta,\mathcal{I}) \int_{\omega^T}^{\overline{\omega}} \xi^{ind}(\theta,\mathcal{I},\omega) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I})) dK_{\mathcal{I}}(\theta) dF(\theta)}{(\beta^{co2}+\beta^{hs})y_m}$ *and ω^T is the skill level associated with the threshold income level y^T and y_m is average income in the top bracket.*

Proof. Mechanical effects by a small reform of τ are given by $\overline{M} = (\beta^{co2}+\beta^{hs})(y_m - y^t)\left(1-\overline{\mathcal{W}}\right)$. Labor supply effects are $\overline{LS} = -\frac{\tau}{1-\tau}\epsilon y_m(\beta^{co2}+\beta^{hs})$. Using $\overline{LS}+\overline{M}+\overline{CG}(\beta^{co2}+\beta^{hs})y_m = 0$ and $\frac{a}{a-1} = \frac{y_m}{y^t}$ gives the result. □

$\overline{CG}$ captures the fiscal externality of the top tax rate through enrollment responses. When $\overline{CG}$ is zero, the formula collapse to the seminal Saez (2001) top tax formula. From the discussion in the theory section and the results from the quantified model in Section 6.1, we expect $\overline{CG}$ to be negative as price (substitution) effects are likely to dominate. As a consequence, optimal top tax rates should be lower. This will confirmed next in the optimal policy simulations, although the effects of $\overline{CG}$ will be shown to be very small.

Quantitative Exercise. We consider the case where the government can set income taxes optimally given education subsidies as in the calibration. In Figure 13(a) optimal marginal tax rates are plotted as a function of yearly gross income. The shape of marginal tax rates is U-shaped, as is typically the case with exogenous education (Diamond 1998, Saez 2001). In order to illustrate the impact of the endogeneity of college enrollment, we also calculate a so called *self-confirming policy equilibrium* (SCPE) as defined in Rothschild and Scheuer (2013,14). Here a government is considered that naively sets policies under the assumption that college enrollment does not react to policies. A SCPE is defined as a situation where the college decisions implied by government's policies coincide with the expectation of the naive government. As can be seen in Figure 13(a), marginal tax rates are slightly higher in this case – the maximum difference is 1.2 percentage points. For top income tax rates, the endogeneity of enrollment plays no significant role.

The share of college graduates in this case is 31.88% – this number should be compared to 35.05% enrollment in the absence of borrowing constraints. Intuitively, the Utilitarian planner uses the tax system to redistribute leading to higher taxes. As the planner cannot use subsidies



(a) Optimal Marginal Tax Rates

(b) Graduation Rates

Figure 13: Optimal Income Taxes for Given Education Subsidies

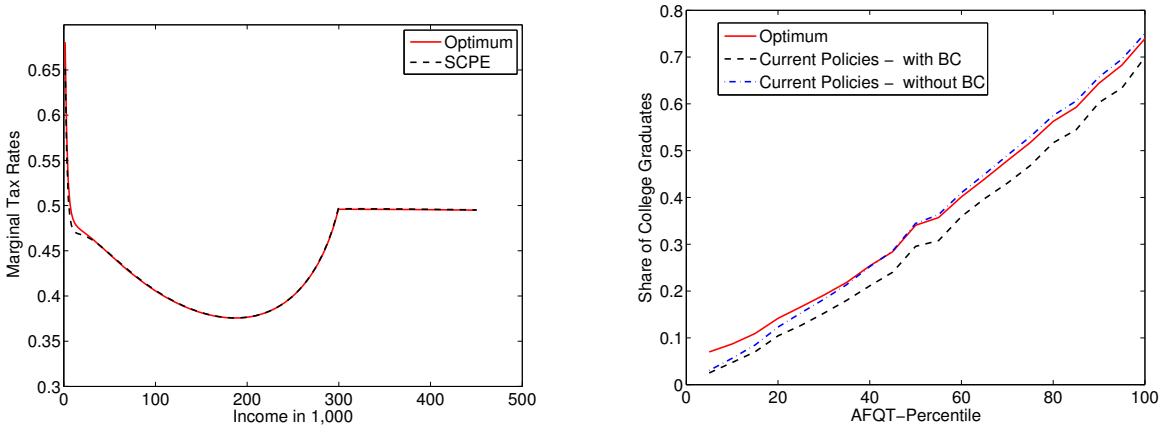
to offset the disincentive effects from labor taxes, the higher level of redistribution necessarily implies a decrease in enrollment. Interestingly, however, enrollment/graduation increases for low ability individuals. The much higher lump sum transfer, which is now at $12,738, implies more insurance against low earning. This is particularly relevant for low AFQT levels, as these individuals are particularly likely to earn low income even with a college degree. The additional amount of insurance allows individuals to fail and therefore encourages them to go to college. This effect is similar as in the theory of the welfare state by Sinn (1995), who argues that insurance achieved by redistribution can promote risky but socially desirable investment if individuals are risk averse.

7.4 Joint Optimal Income Taxes and Education Subsidies

We now explore the case where the government can set income taxes and education subsidies jointly optimal. From a theoretical point of view, optimal subsidies are still characterized by the formula in Proposition 1 and optimal marginal tax rates by the formula in Proposition 3.

Optimal marginal tax rates are displayed in Figure 14(a). They are slightly higher as compared to the previous subsection. This result is intuitive as higher education subsidies can counteract tax distortions and therefore allow for more redistribution and social insurance.

In this scenario, we allow the government also to set college subsidies optimally. The government sets optimal education subsidies to $16,957. This number is much higher than the number in Section 7.2, where we calculated the optimal college subsidy taking the current tax system as given. This is intuitive. Since optimal Utilitarian marginal tax rates are much higher than the ones in the current tax system, the fiscal externality imposed by the decision to go



(a) Optimal Marginal Tax Rates

(b) Graduation Rates

Figure 14: Optimal Income Taxes and College Subsidies

to college are higher which in turn increases the gains from subsidizing college. In addition, higher tax rates allow for more public spending on education.

This also explains why there is not a big difference between optimal marginal tax rates and those in an SCPE. The new term in Proposition 3 is now smaller. Since subsidies are at much higher levels (about 3-4 times the level), the fiscal externality is much smaller.

Lastly, as Figure 14(b) reveals, optimal education subsidies boost college enrollment in particular for lower AFQT-levels far above the numbers in the previous subsection, where subsidies were fixed. Also compared to current policies without borrowing constraints, graduation slightly increases from 35.05% to 35.25%.

7.5 Optimal Education Dependent Taxes

Given that the government can observe the college choice, we now explore how optimal education dependent taxes should look like.

Proposition 4. *Optimal education dependent marginal tax rates satisfy*

$$\frac{\mathcal{T}^{co\prime}(y_{co}(\omega^*))}{1-\mathcal{T}^{co\prime}(y_{co}(\omega^*))} = \left(1+\frac{1}{\varepsilon^{co}(\omega^*)}\right) \times \left(\frac{\int_{\omega^*}^{\overline{\omega}} \beta^{co2}\left(1-\mathcal{W}^{co}(\omega)\right) dG(\omega)}{g^{co}(\omega^*)\beta^{co2}\omega^*} + \frac{\int_{\omega^*}^{\overline{\omega}} \int_{\Theta} \int_{\mathbb{R}_+} \Delta\mathcal{T}(\theta,\mathcal{I})\xi^{co}(\theta,\mathcal{I},\omega) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I})) dK_{\mathcal{I}}(\theta) dF(\theta) d\omega}{g^{co}(\omega^*)\beta^{co2}\omega^*}\right) \tag{30}$$

and

$$\frac{\mathcal{T}^{hs\prime}(y_{hs}(\omega^*))}{1-\mathcal{T}^{hs\prime}(y_{hs}(\omega^*))} = \left(1+\frac{1}{\varepsilon^{hs}(\omega^*)}\right) \times \left(\frac{\int_{\omega^*}^{\overline{\omega}} \beta^{hs}\left(1-\mathcal{W}^{hs}(\omega)\right) dG(\omega)}{g^{hs}(\omega^*)\beta^{hs}\omega^*} + \frac{\int_{\omega^*}^{\overline{\omega}} \int_{\Theta} \int_{\mathbb{R}_+} \Delta\mathcal{T}(\theta,\mathcal{I})\xi^{hs}(\theta,\mathcal{I},\omega) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I})) dK_{\mathcal{I}}(\theta) dF(\theta) d\omega}{g^{hs}(\omega^*)\beta^{hs}\omega^*}\right) \tag{31}$$

Proof. See Appendix A.5.3 □

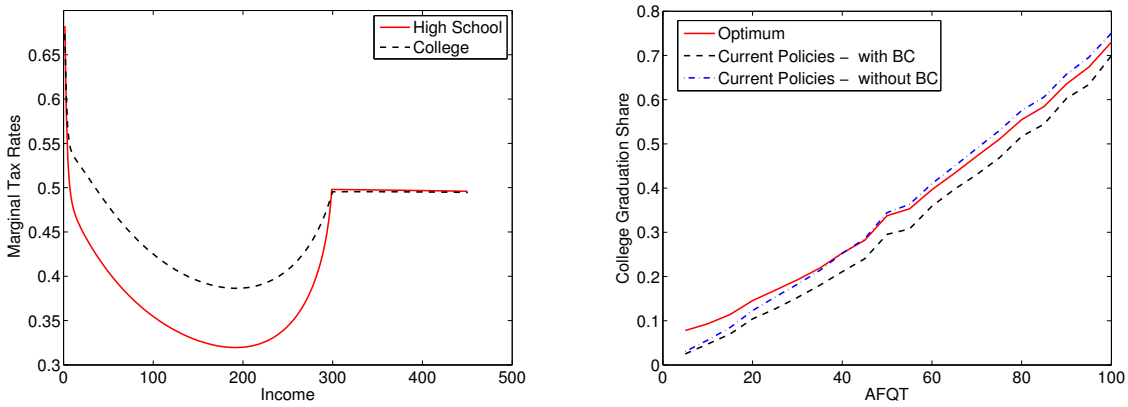
These formulas are very similar to the education independent tax formula in Proposition 3. A difference is that all the terms now carry the superscript *co* or *hs*. It is well known from the tagging literature[39] that conditioning taxes on characteristics other than income is a powerful tool to tailor marginal tax rates to different subgroups. Importantly, education is not an exogenous tag here. Instead it is an endogenous choice. The planner therefore has to take into account the endogeneity of enrollment similarly as in Proposition 3. Whereas for education independent taxes, the effect of higher taxes on enrollment is theoretically ambiguous, this is not the case with education dependent taxes. Thus, for $\Delta\mathcal{T}(\theta) > 0$, endogeneity of enrollment is a clear force towards lower college taxes and higher high school taxes (as compared to the case with exogenous enrollment, i.e. classical tagging). Note that these formulas are related to Scheuer (2013), who considers the differential tax treatment of entrepreneurs and workers.

Quantitative Exercise. Figure 15(a) illustrates marginal tax rates as a function of gross income. Up to income levels of \$350,000 – where the Pareto tail starts for both distributions – marginal tax rates are higher for college graduates. For higher incomes, the result slightly reverses before marginal tax rates start to converge as the enrollment effect approaches zero. The optimal education subsidy is \$23,723. Overall graduation is now at 35.04%. Figure 15(b) illustrates graduation as a function of the AFQT-score.

8 Conclusion

The goal of the paper is to build a transparent and empirically plausible public finance framework to explore the efficient design of college subsidies and income taxation. One of the main innovations is the close connection of the model to the empirical college literature and its successfulness in matching quasi-experimental evidence on individual responses along the college margin.

[39] This literature goes back to Akerlof (1978). For more recent contributions addressing this topic, see Mankiw and Weinzierl (2010), Cremer, Firouz, and Lozachmeur (2010) and Weinzierl (2014).



(a) Optimal Marginal Tax Rates

(b) Graduation Rates

Figure 15: Optimal Education Dependent Taxes

We obtain several analytical and quantitive results for the revenue effects of education and tax policy reforms and optimal education and tax policies. The main results for subsidies are that increasing college subsidies is to a large degree self-financing through higher tax revenue in the future and that targeting different subgroups (e.g. based on parental background) can substantially increase the power of college subsidies and make an increase a free lunch. For income tax design, we find that the excess burden caused by the college margin is rather small. In accordance with this result, the impact of endogenous college graduation on optimal income taxes is rather small.

In future research it would be interesting to consider more levels of college education such as associate degrees, bachelor degrees and master degrees. Differentiating subsidies across college majors is also likely to be a powerful policy instrument that deserves consideration in future research. Application assistance for financial aid might be a cost effective and powerful policy tool because it induces more individuals to go to college (Bettinger, Long, Oreopoulos, and Sanbonmatsu 2012) without increasing the payments made to inframarginal students. It would also be interesting to study the consequences of general equilibrium effects for efficient policy design that we raised in Section 5.4 more deeply.[40] Investigating the interaction of education and taxation policies for other countries than the US is an important task left open for future research. In ongoing research, we use a variation of our model to evaluate whether financial aid targeted to underrepresented minorities (affirmative action) can be grounded on similar efficiency reasons as need-based financial aid.

[40]See Jacobs (2013) for a theoretical investigation of optimal education and tax policies in general equilibrium. Methods developed by Rothschild and Scheuer (2013) are likely to be a good starting point for further research in these directions.

References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2013): "Education Policy and Intergenerational Transfers in Equilibrium," Discussion paper, National Bureau of Economic Research.

Ábrahám, A., S. Koehne, and N. Pavoni (2014): "Optimal income taxation with asset accumulation," *Working Paper*.

Akerlof, G. A. (1978): "The Economics of 'Tagging' as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," *The American Economic Review*, 68(1), 8–19.

Altonji, J., P. Bharadwaj, and F. Lange (2011): "Changes in the characteristics of American youth: Implications for adult outcomes," *Journal of Labor Economics*, 30(4), 783–828.

Anderberg, D. (2009): "Optimal Policy and the Risk Properties of Human Capital Reconsidered," *Journal of Public Economics*, 93(9-10), 1017–1026.

Autor, D. H. (2014): "Skills, education, and the rise of earnings inequality among the "other 99 percent"," *Science*, 344(6186), 843–851.

Badel, A., and M. Huggett (2014): "Taxing top earners: a human capital perspective," *Federal Reserve Bank of St. Louis Working Paper Series*, (2014-017).

Belley, P., and L. J. Lochner (2007): "The changing role of family income and ability in determining educational achievement," Discussion paper, NBER Working Paper 13527.

Best, M., and H. Kleven (2013): "Optimal Income Taxation with Career Effects of Work Effort," *Working Paper, London School of Economics*.

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.

Blundell, R., and A. Shephard (2012): "Employment, Hours of Work and the Optimal Taxation of Low-Income Families," *The Review of Economic Studies*, 79(2), 481–510.

Bohacek, R., and M. Kapicka (2008): "Optimal Human Capital Policies," *Journal of Monetary Economics*, 55(1), 1–16.

Bovenberg, L., and B. Jacobs (2005): "Redistribution and Education Subsidies are Siamese Twins," *Journal of Public Economics*, 89(11-12), 2005–2035.

Card, D. (1999): "The causal effect of education on earnings," *Handbook of labor economics*, 3, 1801–1863.

Carneiro, P., and J. Heckman (2005): " Human Capital Policies," in *Inequality in America: What Role for Human Capital Policies*, pp. 77–240. The MIT Press.

Castleman, B. L., and B. T. Long (2013): "Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation," *NBER Working Paper 19306*.

Chatterjee, S., and F. Ionescu (2012): "Insuring student loans against the financial risk of failing to complete college," *Quantitative Economics*, 3(3), 393–420.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review Papers and Proceedings*, 101, 471–475.

Cremer, H., G. Firouz, and J.-M. Lozachmeur (2010): "Tagging and Income Taxation: Theory and an Application," *American Economic Journal: Economic Policy*, 2(1).

Cunha, F., J. Heckman, and S. Navarro (2005): "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economic Papers*, 57(2), 191–261.

Cunha, F., F. Karahan, and I. Soares (2011): "Returns to skills and the college premium," *Journal of Money, Credit and Banking*, 43(1), 39–86.

Deming, D., and S. Dynarski (2009): "Into college, out of poverty? Policies to increase the postsecondary attainment of the poor," *NBER Working Paper 15387*.

Diamond, P. A. (1998): "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 88(1), 83–95.

Diamond, P. A., and E. Saez (2011): "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, 25(4), 165–90.

Findeisen, S., and D. Sachs (2013): "Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans," *Working Paper*.

Findeisen, S., and D. Sachs (2014): "Efficient Labor and Capital Income Taxation over the Life Cycle," *Working Paper*.

Goldin, C. D., and L. F. Katz (2009): *The race between education and technology*. Harvard University Press.

Golosov, M., A. Tsyvinski, and N. Werquin (2014): "A Variational Approach to the Analysis of Tax Systems," *Working Paper*.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, capacity utilization, and the real business cycle," *The American Economic Review*, pp. 402–417.

Gruber, J., and E. Saez (2002): "The elasticity of taxable income: evidence and implications," *Journal of Public Economics*, 84(1), 1–32.

Guner, N., R. Kaygusuz, and G. Ventura (2013): "Income Taxation of U.S. Households: Facts and Parametric Estimates," *Review of Economic Dynamics (forthcoming)*.

Hamilton, J. H. (1987): "Optimal Wage and Income Taxation with Wage Uncertainty," *International Economic Review*, 28(2), 373–388.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006): "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond," *Handbook of the Economics of Education*, 1, 307–458.

Hilger, N. G. (2014): "Parental credit constraints and college attendance: Evidence from timing of parental layoffs," *Working Paper*.

Holter, H. A. (2014): "Accounting for cross-country differences in intergenerational earnings persistence: the impact of taxation and public education expenditure," .

Jacobs, B. (2013): "Optimal redistributive tax and education policies in general equilibrium," *International Tax and Public Finance*, 20(2), 312–337.

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.

Jacobs, B., D. Schindler, and H. Yang (2012): "Optimal Taxation of Risky Human Capital," *Scandinavian Journal of Economics*, 114(3), 908–931.

Jacquet, L., E. Lehmann, and B. Van der Linden (2013): "The Optimal Marginal Tax Rates with both Extensive and Intensive Responses," *Journal of Economic Theory*, 148(5), 1770–1805.

Johnson, M. T. (2013): "Borrowing Constraints, College Enrollment, and Delayed Entry," *Journal of Labor Economics*, 31(4), 669–725.

Kane, T. J. (2006): "Public intervention in post-secondary education," *Handbook of the Economics of Education*, 2, 1369–1401.

Kapicka, M. (2013): "Optimal Mirrleesean Taxation with Unobservable Human Capital Formation," *Working paper, UC Santa Barbara.*

Kapicka, M., and J. Neira (2013): "Optimal taxation in a life-cycle economy with endogenous human capital formation," *Working paper, UC Santa Barbara.*

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

Kleven, H. J., and E. A. Schultz (2012): "Estimating Taxable Income Responses using Danish Tax Reforms," .

Krueger, D., and A. Ludwig (2013a): "Optimal Capital and Progressive Labor Income Taxation with Endogenous Schooling Decisions and Intergenerational Transfers," *Working Paper*, 103(3), 496–501.

——— (2013b): "Optimal Progressive Labor Income Taxation and Education Subsidies when Education Decisions and Intergenerational Transfers are Endogenous," *American Economic Review Papers & Proceedings*, 103(3), 496–501.

Laun, T., and J. Wallenius (2013): "Social Insurance and Retirement: A Cross-Country Perspective," *Working Paper, Uppsala University.*

Lawson, N. (2014): "Optimal College Tuition Subsidies," *Working Paper.*

Lee, D., S. Y. Lee, and Y. Shin (2014): "The Option Value of Human Capital," *Working Paper.*

Lehmann, E., L. Simula, and A. Trannoy (2014): "Tax Me if You Can: Optimal Income Tax Competition between Competing Governments," *Quarterly Journal of Economics (forthcoming).*

Lochner, L. J., and A. Monge-Naranjo (2011): "The Nature of Credit Constraints and Human Capital," *The American Economic Review*, 101(6), 2487–2529.

Lochner, L. J., and A. Monge-Naranjo (2012): "Credit Constraints in Education," *Annual Review of Economics*, 4(1), 225–256.

Lovenheim, M. F. (2011): "The Effect of Liquid Housing Wealth on College Enrollment," *Journal of Labor Economics*, 29(4), pp. 741–771.

Mankiw, N., and M. Weinzierl (2010): "The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution," *American Economic Journal: Economic Policy*, 2(1), 155–176.

Manoli, D., and N. Turner (2014): "Cash-on-Hand & College Enrollment: Evidence from Population Tax Data and Policy Nonlinearities," *NBER Working Paper No. 19836*.

Mendoza, E. G., and V. Z. Yue (2012): "A general equilibrium model of sovereign default and business cycles," *The Quarterly Journal of Economics*, 127(2), 889–946.

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *The Review of Economic Studies*, 38(2), 175–208.

Navarro, S. (2011): "Using observed choices to infer agent's information: Reconsidering the importance of borrowing constraints, uncertainty and preferences in college attendance," *CIBC Working Paper*.

OECD (2014): "Education at a Glance 2011: OECD Indicators, OECD Publishing," URL: http://dx.doi.org/10.1787/eag-2014-en.

Okun, A. M. (1975): *Equality and efficiency, the big tradeoff*. Brookings Institution Press.

Oreopoulos, P., and U. Petronijevic (2013): "Making college worth it: A review of research on the returns to higher education," *NBER Working Paper 19053*.

Piketty, T. (1997): "La redistribution fiscale face au chomage," *Revue française d'économie*, 12(1), 157–201.

Piketty, T., and E. Saez (2013): "Optimal labor income taxation," *Handbook of Public Economics, Vol 5*.

Richter, W. F. (2009): "Taxing education in Ramsey's tradition," *Journal of Public Economics*, 93(11), 1254–1260.

Rothschild, C., and F. Scheuer (2013): "Redistributive Taxation in the Roy Model," *The Quarterly Journal of Economics*, 128(2), 623–668.

——— (2014): "Optimal Taxation with Rent-Seeking," *Working Paper*.

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68(1), 205–229.

Saez, E. (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117(3), 1039–1073.

Scheuer, F. (2014): "Entrepreneurial Taxation with Endogenous Entry," *American Economic Journal: Economic Policy*, 6, 126–163.

Schindler, D., and H. Yang (2014): "Catalysts for social insurance: education subsidies versus physical capital taxation," *International Tax and Public Finance (forthcoming)*.

Shapiro, D., A. Dundar, J. Chen, M. Ziskin, E. Park, V. Torres, and Y.-C. Chiang (2012): "Completing College: A National View of Student Attainment Rates. Signature [TM] Report 4.," *National Student Clearinghouse*.

Sinn, H.-W. (1995): "A Theory of the Welfare State," *Scandinavian Journal of Economics*, 97(4), 495–526.

Snyder, T. D., and S. A. Dillow (2013): *Digest of education statistics 2012*. National Center for Education Statistics.

Snyder, T. D., and C. M. Hoffman (2001): *Digest of education statistics 2000*. National Center for Education Statistics.

Stantcheva, S. (2014): "Optimal Taxation and Human Capital Policies Over the Lifecycle," *Working Paper, Harvard University*.

Stinebrickner, R., and T. R. Stinebrickner (2003): "Understanding Educational Outcomes of Students from Low-Income Families Evidence from a Liberal Arts College with a Full Tuition Subsidy Program," *Journal of Human Resources*, 38(3), 591–617.

Varian, H. R. (1980): "Redistributive taxation as social insurance," *Journal of Public Economics*, 14(1), 49–68.

Weinzierl, M. (2014): "The Promise of Positive Optimal Taxation," *Journal of Public Economics (forthcoming)*.

Winter, C. (2014): "Accounting for the Changing Role of Family Income in Determining College Entry," *Scandinavian Journal of Economics*, 116(4), 909–963.

Zimmerman, S. (2014): "The returns to college admission for academically marginal students," *Journal of Labor Economics*, 32(4), 711–754.

A Appendix

A.1 Details on Income Regressions

For most individuals, we do not have information in every year. First of all, we never have information after age 53. Second, since 1994 the survey is conducted biannually. Third, we often have to deal with missing values. To resolve the first issue, we assume that incomes are

flat afterwards, which is roughly what one finds in data sets with information on earnings over the whole life cycle. See, e.g., Figures 13 and 14 in Lee, Lee, and Shin (2014). Concerning the second issue, we take the average of the income in the year before and after. Concerning the third issue, we proceed similarly but also take values that are two and three years away if information for the year before and after is missing as well. All other years that are still missing are then just not taken into account for calculating $\tilde{y}_i$. Assume, e.g., that only income at age 19, 33 and 46 were observable. Then we would calculate

$$\tilde{y}_i = \frac{\frac{y_{i19}}{(1+r)} + \frac{y_{i33}}{(1+r)^{14}} + \frac{y_{i46}}{(1+r)^{27}}}{\frac{1}{(1+r)} + \frac{1}{(1+r)^{14}} + \frac{1}{(1+r)^{27}}}.$$

All incomes are measured in 2000 dollars.

Our estimates for the slopes are. $\hat{\beta}_{co}^{IN} = 1.88$ (0.186) and $\hat{\beta}_{hs}^{IN} = 1.28$ (0.074). As described in the main text, the second-moments of the log-normal parts are education dependent, so that until 350k, $\ln y$ is normal with standard deviation σ^e. We directly take the estimates for σ^e from the distribution of residuals from (6). The values are 0.6548 for college and 0.6631 for high-school.

A.2 Tuition Fees and Public Costs of Colleges

First, we categorize the following 4 regions:

- Northeast: CT, ME, MA, NH, NJ, NY, PA, RI, VT
- North Central: IL, IN, IA, KS, MI, MN, MO, NE, OH, ND, SD, WI
- South: AL, AR, DE, DC, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN , TX, VA, WV
- West: AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, WY

We base the following calculations on numbers presented by Snyder and Hoffman (2001). Table 313 of this report contains average tuition fees for four-year public and private universities. According to Table 173, 65% of all four-year college students went to public institutions, whereas 35% went to private institutions. For each state we can therefore calculate the average (weighted by the shares) tuition fee for a four-year college. We then use these numbers to calculate the average for each of the four regions, where we weigh the different states by their population size. We then arrive at numbers for yearly tuition & fees of \$9,435 (Northeast), \$7,646 (North Central), \$6,414 (South) and \$7,073 (West). For all individuals in the data with missing information about their state of residence, we chose a country wide population size weighted average of \$7,434.

Tuition revenue of colleges typically only covers a certain share of their expenditure. Figures 18 and 19 in Snyder and Hoffman (2001) illustrate by which sources public and private colleges finance cover their costs. Unfortunately no distinction between two and four-year colleges is available. From Figures 18 and 19 we then infer how many dollars of public appropriations are spent for each dollar of tuition. Many of these public appropriations are also used to finance graduate students. It is unlikely that the marginal public appropriation for a bachelor student therefore equals the average public appropriation at a college given that costs for graduate students are higher. To solve this issue, we focus on institutions "that primarily focus on undergraduate education" as defined in Table 345. Lastly, to avoid double counting of grants and fee waivers, we exclude them from the calculation as we directly use the detailed individual data about financial aid receipt from the NLSY (see Section 4.5). Based on these calculations we arrive at marginal public appropriations of \$5,485 (Northeast), \$4,514 (North Central), \$3,558 (South), \$3,604 (West) and \$4,157 (No information about region).

A.3 Need-Based Grant Reforms and Parental Incentives

Assume that parents have preferences of the form:

$$u(c, K, \mathcal{I}),$$

where K is total resources available for their kids. For the child's resources, we have $K = tr_{co} + \mathcal{G}(\theta, \mathcal{I})$. Parents choose their transfer tr_{hs}. The budget constraint of the parents is $c = \mathcal{I} - T(\mathcal{I}) - tr_{co}$. Their problem then reads as:

$$\max_{tr_{co}, \mathcal{I}} u\left(y - T(\mathcal{I}) - tr_{co}, tr_{co} + \mathcal{G}(\theta, \mathcal{I}), \mathcal{I}\right)).$$

The first-order conditions are

$$u_K = u_c$$

and

$$u_c \cdot [1 - T'(\mathcal{I})] + u_K \cdot \left[\frac{\partial \mathcal{G}(\theta, \mathcal{I})}{\partial \mathcal{I}}\right] = -u_{\mathcal{I}}$$

Combining the first-order conditions yields:

$$u_c \cdot \left[1 - T'(\mathcal{I}) + \frac{\partial \mathcal{G}(\theta, \mathcal{I})}{\partial \mathcal{I}}\right] = -u_{\mathcal{I}}$$

which implies that the parents will react to an decrease in $\frac{\partial \mathcal{G}(\theta, \mathcal{I})}{\partial \mathcal{I}}$ in the same way as to an increase in $\mathcal{T}'(\mathcal{I})$. Denote by $\varepsilon_{\mathcal{I}^*, 1 - T' + \frac{\partial \mathcal{G}(\theta, \mathcal{I})}{\partial \mathcal{I}}}$ the elasticity of income with respect to an increase of the effective marginal tax rate.

The reform we are considering is to increase grants for children with parental income below a certain threshold $\mathcal{I}^*$ by $d\mathcal{G}$. This implies an increase of $\frac{\partial \mathcal{G}(\theta,\mathcal{I})}{\partial \mathcal{I}}$ within a small interval $[\mathcal{I}^*, \mathcal{I}^* + \Delta\mathcal{I}]$ by

$$\Delta\left(\frac{\partial \mathcal{G}(\theta,\mathcal{I})}{\partial \mathcal{I}}\right) = \frac{d\mathcal{G}}{\Delta\mathcal{I}}$$

.

The impact of the change in parents behavior on public funds is then given by:

$$\left(\mathcal{T}' - \frac{\partial \mathcal{G}(\theta,\mathcal{I})}{\partial \mathcal{I}}\right) \varepsilon_{\mathcal{I}^*, 1-\mathcal{T}'+\frac{\partial \mathcal{G}(\theta,\mathcal{I})}{\partial \mathcal{I}}} \frac{\mathcal{I}^*}{1-\mathcal{T}'+\frac{\partial \mathcal{G}(\theta,\mathcal{I})}{\partial \mathcal{I}}} \frac{d\mathcal{G}}{\Delta\mathcal{I}}.$$

The mass of affected parents is $h^{co}_{\mathcal{I}^*}\Delta\mathcal{I}$,where $h^{co}_{\mathcal{I}^*}$ is the density of parents with income $\mathcal{I}^*$ and children that go to college. The overall effect is therefore

$$\left(\mathcal{T}' - \frac{\partial \mathcal{G}(\theta,\mathcal{I})}{\partial \mathcal{I}}\right) \varepsilon_{\mathcal{I}^*, 1-\mathcal{T}'+\frac{\partial \mathcal{G}(\theta,\mathcal{I})}{\partial \mathcal{I}}} \frac{\mathcal{I}^*}{1-\mathcal{T}'+\frac{\partial \mathcal{G}(\theta,\mathcal{I})}{\partial \mathcal{I}}} d\mathcal{G}.$$

A.4 Revenue Effects of Lump-Sum Tax Reforms

Consider a small increase of the lump-sum element of the tax function for all individuals by $d\mathcal{T}(0)$. This mechanically increases tax revenue in NPV by

$$M(0) = d\mathcal{T}(0)\left(\beta^{co2}G^{co}(\overline{\omega}) + \beta^{hs}G^{hs}(\overline{\omega})\right).$$

The increase in the lump sum element of the tax function also changes incentives for enrollment. In fact, graduation will increase by:

$$d\mathcal{T}(0)\int_{\Theta}\int_{\mathbb{R}_+}\int_{\Omega} \xi^{ind}(\theta,\mathcal{I},\omega)h_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))dK_{\mathcal{I}}(\theta)dF(\theta).$$

This implies a first-order effect on the government budget which is given by:

$$CG(0) = d\mathcal{T}(0)\int_{\Theta}\Delta\mathcal{T}(\theta)\int_{\mathbb{R}_+}\int_{\Omega} \xi^{ind}(\theta,\mathcal{I},\omega)H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta,\mathcal{I}))dK_{\mathcal{I}}(\theta)dF(\theta). \tag{32}$$

The overall impact on public funds of a marginal increase of the lump-sum tax by $\Delta\mathcal{T}(0)$ is thus given by:

$$\Gamma(0) = M(0) + CG(0). \tag{33}$$

We also define a normalized revenue gain measure that is simpler to interpret

$$\gamma(0) = \frac{\Gamma(0)}{d\mathcal{T}(0)\left(\beta^{co}G^{co}(\overline{\omega}) + \beta^{hs}G^{hs}(\overline{\omega})\right)}.$$

This term reads as follows: For each dollar the government mechanically raises by increasing $\mathcal{T}(0)$, it keeps $\gamma(0)$ dollars. Whether this number is larger or smaller than unity solely depends on the sign of (32), which in turn depends on (i) whether higher lump sum taxes lead to higher graduation rates and (ii) whether the average marginal college graduate provides a larger share to the government budget as compared to the case where she had not gone to college. There is not sufficient empirical evidence on the graduation responses with respect to lump sum taxes – or welfare benefits – that would allow us to calculate $\gamma(0)$ in a sufficient-statistic manner. Using our parameterized model, we find that an increase in lump-sum taxes increases graduation and that these marginal students increase the net tax revenue of the government leading to a number $\gamma(0) = 1.04$. For each marginal dollar the government raises with increasing lump sum taxes, it obtains another \$0.04 due to increased college graduation. Intuitively, higher lump-sum taxes – or lower welfare benefits – make being poor less attractive and therefore rise incentives to invest in education by obtaining a college degree. For a borrowing limit of \$32,250 and a risk aversion parameter of 1, $\gamma(0)$ gets as high as 1.07.

Education-Dependent Lump-Sum Reforms. In case that lump-sum taxes are only increased for one education group, the mechanical increase in tax revenue is:

$$\forall\, i = hs, co: \;\; M^i(0) = d\mathcal{T}^i(0)\beta^i G^i(\overline{\omega}).$$

The college graduation effect reads as

$$\forall\, i = hs, co: \;\; CG^i(0) = d\mathcal{T}^i(0) \int_{\Theta} \Delta\mathcal{T}(\theta) \int_{\mathbb{R}_+} \int_{\Omega} \xi^i(\theta, \mathcal{I}, \omega) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_{\mathcal{I}}(\theta) dF(\theta).$$

Based on that, we can also define the overall impact on welfare

$$\forall\, i = hs, co: \;\; \Gamma^i(0) = M^i(0) + CG^i(0).$$

The normalized version reads as

$$\gamma^i(0) = \frac{\Gamma^i(0)}{d\mathcal{T}^i(0)\beta^i G^i(\overline{\omega})}.$$

Again, there are no well-established empirical results that would help us to evaluate this equation directly. Using the quantitative version of our model from Section 4, we obtain $\gamma^{hs}(0) = 1.15$ and $\gamma_R^{co}(0) = 0.81$. Raising lump sum elements in an education dependent fashion clearly acts like a price subsidy. Increasing it solely for college graduates decreases the number of college graduates and the government loses \$0.19 for each dollar it mechanically

raises.[41] For high school lump sum taxes, this effect reverses. Higher lump sum taxes for high school graduates make going to college more attractive and therefore the government obtains an additional \$0.15 for each mechanically raised dollar. For $\gamma = 1$ and a borrowing limit of \$32,250, effects get much stronger: $\gamma_R^{co}(0) = 0.57$ and $\gamma_R^{hs}(0) = 1.26$.

A.5 Proofs Section 7

A.5.1 Proof of Proposition 1

Formula (9) in Section 5.1 gives the impact on government revenue for a marginal increase in grants. Inserting (7) and (8) yields

$$\Gamma^{\mathcal{G}} = d\mathcal{G} \times \int_{\Theta} \int_{\mathbb{R}_+} \Delta\mathcal{T}(\theta, \mathcal{I}) \xi^{\mathcal{G}}(\theta, \mathcal{I}) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_{\theta}(\mathcal{I}) dF(\theta) + d\mathcal{G} \times \beta^{co1} G^{co}(\overline{\omega}).$$

This expression does not take into account the impact on utility of students as the focus was solely on revenue. Taking into account utility as well and using the definition of the social welfare weights of students as defined in (26) changes the expression to

$$d\mathcal{G} \times \int_{\Theta} \int_{\mathbb{R}_+} \Delta\mathcal{T}(\theta, \mathcal{I}) \xi^{\mathcal{G}}(\theta, \mathcal{I}) H_{\theta,\mathcal{I}}(\tilde{\kappa}(\theta, \mathcal{I})) dK_{\theta}(\mathcal{I}) dF(\theta) + d\mathcal{G} \times \beta^{co1} G^{co}(\overline{\omega}) \left(1 - \mathcal{W}^{st}\right).$$

Setting this expression to zero and using the standard definition of an elasticity gives the formula in Proposition 1.

A.5.2 Proof of Proposition 3

Equation (18) gives the revenue gain of a marginal increase in the marginal tax rate at income level $y(\omega^*)$. The effect of individual utility is not taken into account. Taking it also into account the expression reads as

$$d\mathcal{T}' dy \int_{\omega^*}^{\overline{\omega}} \left(\beta^{co2} \left(G^{co}(\overline{\omega}) - G^{co}(\omega^*)\right) \left(1 - \mathcal{W}^{co}(\omega)\right) + \beta^{hs} \left(G^{hs}(\overline{\omega}) - G^{hs}(\omega^*)\right)\right) \left(1 - \mathcal{W}^{hs}(\omega)\right) dF(\omega)$$
$$+ CG(y(\omega^*)) + LS(y(\omega^*)). \tag{34}$$

For a tax function to be optimal, this has to be zero. Setting the term to zero and solving for $\frac{\mathcal{T}'(y(\omega))}{1-\mathcal{T}'(y(\omega))}$ yields Proposition 3.

[41] Note that this result is still much weaker than the result from Section 5 where a grant increase was even self-financing. Based on that result, one could suspect that an increase of the lump sum tax solely for college graduates would actually decrease tax revenue as the decrease in graduation is that large. The simple reason for this asymmetry of results are borrowing constraints. As individuals are borrowing constrained, one dollar more of consumption during college has a much stronger effect than one more dollar as a worker. See Lemma 2.

A.5.3 Proof of Proposition 4

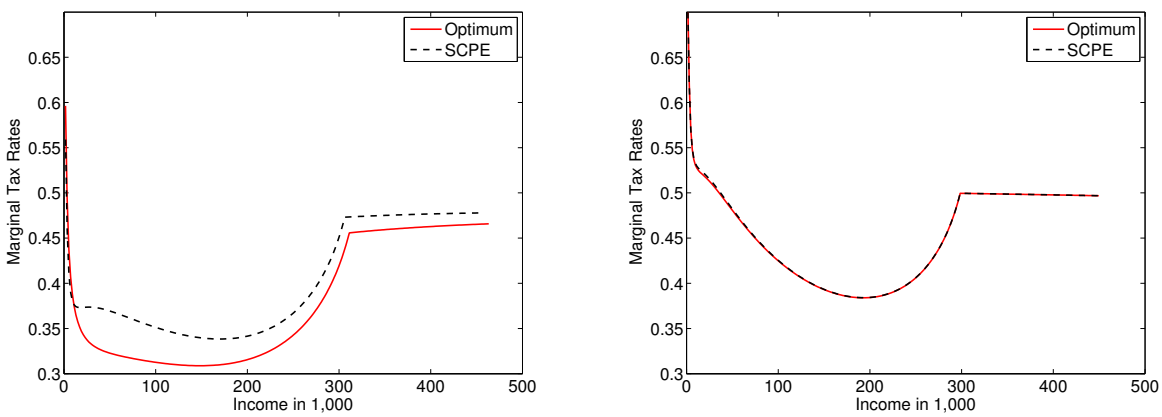
Equation (21) gives the revenue gain of a marginal increase in the marginal tax rate at income level $y_{co}(\omega^*)$ or $y_{hs}(\omega^*)$. The effect of individual utility is not taken into account. Taking it also into account the expression reads as (for $i = co, hs$)

$$d\mathcal{T}^{i'} dy \beta^i \int_{\omega^*}^{\omega} \beta^i \left(G^i(\overline{\omega}) - G^i(\omega^*)\right) \left(1 - \mathcal{W}^i(\omega)\right) dF(\omega) + CG^i(y^i(\omega^*)) + LS^i(y^i(\omega^*)).$$

For a tax function to be optimal, this has to be zero. Setting the term to zero and solving for $\frac{\mathcal{T}^{i'}(y(\omega))}{1-\mathcal{T}^{i'}(y(\omega))}$ yields the formulas in Proposition 4.

A.6 Robustness Mirrleesian Income Taxation

Figure 16(a) shows that differences in marginal tax rates can get more substantial for lower values of the CRRA coefficient. The difference in optimal marginal tax rates and that of a naive social planner differ up to 4.7 percentage points. By contrast, if the CRRA coefficient is higher than our benchmark, the difference gets negligible as Figure 16(b) reveals. These figures refer to the case where college subsidies are not set optimally but rather to their current level. If one looks at the case, where subsidies are optimal, the difference also becomes negligible for the case of a CRRA of 1, see Figure 17. Thus, in most cases the formula for optimal marginal tax rates of a planner that does not take into account the college margin are very similar to optimal marginal tax rates.



(a) CRRA=1 (b) CRRA=3

Figure 16: Robustness Optimal Income Taxes

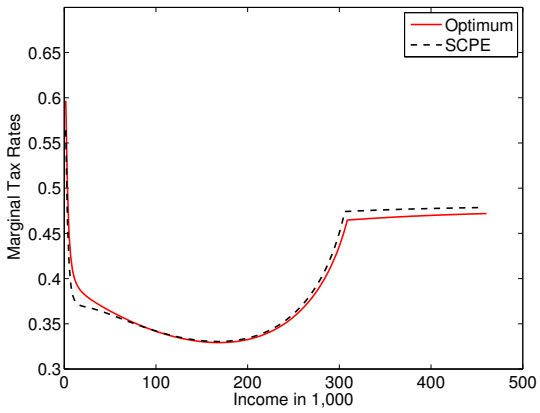


Figure 17: Robustness Optimal Income Taxes with Optimal Subsidies: CRRA=1