# Optimal Need-Based Financial Aid*

**Sebastian Findeisen**
University of Mannheim

**Dominik Sachs**
European University Institute

June 2, 2016

**Abstract**

We study the optimal design of student financial aid and find that it should be declining in parental income. This progressivity result is based on efficiency considerations and does not hinge on a redistributive motive. The key force is the increasing share of inframarginal students. It is robust to different assumptions on credit markets, merit based elements, and income taxes (optimal or current US). A larger degree of progressivity can even be implemented in a Pareto improving way: an increase in financial aid for low parental income children is likely to be self-financing through higher future tax revenue.

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

**Keywords:** Financial Aid, College Subsidies, Optimal Taxation, Inequality

*Contact: findeisen@uni-mannheim.de, dominik.Sachs@eui.eu. An older version of this paper was called *Designing Efficient College and Tax Policies*. We thank Rüdiger Bachmann, Felix Bierbrauer, Richard Blundell, Christian Bredemeier, Friedrich Breyer, David Card, Pedro Carneiro, Alexander Gelber, Marcel Gerard, Emanuel Hansen, Nathan Hendren, Bas Jacobs, Leo Kaas, Marek Kapicka, Kory Kroft, Fabian Krueger, Paul Klein, Tim Lee, Lance Lochner, Normann Lorenz, Thorsten Louis, Alex Ludwig, Marti Mestieri, Michael Peters, Emmanuel Saez, Aleh Tsyvinski, Gianluca Violante, Matthew Weinzierl, Nicolas Werquin and seminar participants at Berkeley, Bocconi, Bonn (MPI & MEE), CEMFI, Dortmund, EIEF, EUI, Frankfurt, IFS/UCL, Louvain (CORE), Notre Dame, Lausanne, Queen's, Salzburg, St. Gallen, Toronto, Toulouse, Uppsala, Warwick as well as conference participants at CEPR Public Meeting 16, CESifo Public Sector Economics, IIPF, NBER SI Macro Public Finance, NORMAC, SEEK and the Taxation Theory Conference. 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.

# 1 Introduction

Most governments provide significant financial aid to college students. At one extreme, in Scandinavian countries college students pay low or no tuition fees and in addition receive grants from the government. In the US, students bear a larger burden of tuition costs; still federal and state spending on grants to college students exceeded 55 billion dollars for the academic year 2014-2015. Most countries additionally target low income students with their policies. In the US, federal spending on the Pell Grant program for low income students exceeded 30 billion in 2014-2015 and has grown by over 80% in the last 10 years (College Board 2015). Despite their importance and potential implications for growth, income inequality, and social mobility, there is surprisingly little research on the normative side on financial aid policies.[1] The contribution of this paper is to explore the optimal design of financial aid policies and clarify the main underlying trade-offs theoretically and quantitatively in a very transparent manner.

We begin by characterizing optimal financial aid policies theoretically. We keep the model as general as possible at this stage without imposing many restrictions on the underlying heterogeneity in the population and express our optimality conditions in terms of sufficient statistics (Chetty 2009).[2] We show that the forces for optimal education subsidies along the *extensive* college margin are quite distinct compared to the case of continuous educational investment[3] (Bovenberg and Jacobs 2005, Stantcheva 2016) and mainly depend on three sufficient statistics: optimal financial aid increases in the share of marginal students and the fiscal externality per marginal student, which captures the marginal benefit of the subsidy. Optimal financial aid decreases in the share of inframarginal students which captures the marginal costs. These three sufficient statistics vary quite a lot among different subgroups of the population which implies that 'tagging' along the extensive margin can be a powerful tool.[4] The form of tagging we then study is to condition on parental income, i.e. we study the optimal need-based student financial aid. We also study the case where the government can optimally set financial aid jointly on the combination of parental income and academic merit.

We implement our model empirically to characterize the desirability of progressive financial aid policies in the US. To this end, we estimate and calibrate the structural parameters of

[1] Recent exceptions are Krueger and Ludwig (2013, 2016) and Lawson (2016). They do not study the need-based aspect of financial aid, however.

[2] The sufficient statistic approach has been applied extensively in the optimal social insurance literature and optimal taxation, (Chetty and Finkelstein 2013, Piketty and Saez 2013). Maybe surprisingly, it has not been used a lot to characterize education and financial aid policies, although governments are heavily involved in the subsidization of education and the stakes in terms of governments' budgets are large.

[3] This resembles the results in the optimal tax literature, where the underlying trade-offs are quite different if labor supply is modelled along the extensive instead of the intensive margin (Diamond 1980, Saez 2002).

[4] 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 model with heterogeneity in parental income, ability and preferences for college, stressing the importance and interaction of parental income and ability to explain the selection into college and heterogenous returns to college graduation. We validate our approach by replicating quasi-experiments on the effects of student financial aid expansions and parental income on enrollment. Our model also predicts returns for marginal students similar to existing estimates from quasi-experiments. The structural nature of the model is nevertheless essential to predict values for the sufficient statistics if we move away from current towards optimal policies.

We have two main results. First, we find that optimal financial aid policies are strongly progressive. In our preferred specification, the level of financial aid drops by more than 60% moving from 5th percentile of the parental income distribution to the 95th percentile. The strong progressivity is very robust and holds for for a broad range of different parameter choices: different tax functions, welfare criteria (Utilitarian versus tax revenue maximizing), and assumptions on credit markets. Second, our estimates suggest that targeted increases in financial aid for low-income students, approximately between the 15th and 45th percentile of the parental income distribution, are self-financing by increases in future tax-revenue, implying that targeted financial aid expansions could be Pareto improving free-lunch policies. Both results point out that financial aid policies for students are a rare case where there is no equity-efficiency trade-off: education policies which lead to a cost effective distribution of aid to help students pay for the cost of college are also in line with redistributive concerns and social mobility.

The paper has three major parts. In the first step in Section 2, we build the framework and characterize how financial aid should optimally vary with parental income. The marginal gain from increasing financial aid for a given parental income level is proportional to the *fiscal externality* it creates. The fiscal externality results from the increase in future tax revenue, which is determined by the returns of college attendance for marginal students as well as the mass of marginal students who are induced to attend college. The latter number has been estimated in numerous studies for different policy reforms in the US (see Castleman and Long (2016) for a recent contribution or Deming and Dynarski (2009) for a survey). The marginal cost of increasing financial aid for a given parental income level is proportional to the amount of inframarginal students – those who would attend even in the absence of the reform. The marginal cost is further scaled down by the welfare weight placed on students from the parental income group.[5] Equating marginal costs and benefits yields an easy to interpret formula for the optimal financial aid level at each parental income level.

[5] The marginal cost is multiplied by $(1-x)$ where $x$ is the marginal social welfare weight placed of the respective group of students (Saez and Stantcheva 2016).

In the second part, in our empirical analysis we specify preferences and the underlying heterogeneity of the model and quantify it in Section 4. This is an essential step since not all elasticities needed to evaluate the optimal policy formulas have been estimated in the previous literature and in order to consider optimal policies we also need to go beyond local responses estimated around current policies. We use data from the National Longitudinal Survey of Youth 1979 and 1997 (henceforth, NLSY79 and NLSY97). Our empirical approach focuses on heterogeneity in three dimensions and their correlations: parental income, ability determined before college and preferences for college attendance. Parental income matters because it is strongly linked to parental transfers during college; heterogeneity in ability is important because it allows for heterogenous returns to college attendance. In particular, it is plausible that marginal students have different returns than inframarginal students. Moreover these returns should differ across the parental income distribution. We estimate the joint distributions of parental income, ability determined before college and preferences for college attendance and their mapping into returns to college and parental transfers. All this pins down the relevant parameters/sufficient statistics (which are not policy invariant) as described in the theoretical part. Our quantitative model can replicate key patterns on how college education varies with parental income and a measure of ability for young adults. Further, our model yields (marginal) returns to college that are in line with the empirical literature (Card 1999, Oreopoulos and Petronijevic 2013, Zimmermann 2014) and can replicate quasi-experimental studies.[6]

The quantitive analysis yields the result that the optimal subsidy is strictly decreasing in parental income. The result is surprisingly robust to the social welfare function, the existence or non-existence of borrowing constraints, and other parameters. In particular, we find that even for a government purely interested in maximizing tax revenue, progressive financial aid is the best policy. One may have expected that fiscal returns to financial aid programs are higher for higher parental income levels, as those children are thought of as being better prepared for college and having higher returns. The latter will indeed be true in our model. However, this effect is clearly dominated by the fact the at higher income level much more students are inframarginal. As a result, optimal subsidies are progressive and the fiscal returns to financial aid expansions are significantly higher for low income children.

[6] A number we target is 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 1.5 percentage points. A number that is the average in 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 0.08 percentage points increase in college graduation. A number that is in line – though slightly smaller – with Hilger (2015). The latter number was not a target for our calibration.

The third major part of the article also takes into account the optimal design of income taxes. Despite the large underlying degree of heterogeneity in the model, we can solve for the fully optimal schedule in the spirit of Mirrlees (1971). First, this exercise is motivated by the fact that higher and more progressive taxes are a complement to financial aid. It is hence important to know how endogenously chosen optimal taxation affects financial aid policies. Second, it allows the government to directly tackle redistributional concerns by progressive taxation instead of redistributing through progressive financial aid. We theoretically characterize optimal tax policies in Section 3 and quantitatively in Section 6. The main result is that optimal financial aid policies are unchanged compared to the case with exogenous taxes: the optimal system features high progessiveness and a high negative dependence on parental income. Although optimal taxes are significantly higher and more progressive than the current system, the main result is not overturned. The intuition here is again that the fraction of inframarginal students is significantly higher for high-income children for basically any tax system.

This paper contributes to the literature studying the optimal design of human capital policies. This has been done in different contexts. Stantcheva (2016) characterizes optimal history-dependent tax and human capital policies in a dynamic life-cycle model. Bovenberg and Jacobs (2005) consider a static model with a continuous education choice and emphasize that education subsidies and taxes are complements, calling them 'siamese twins'. In Findeisen and Sachs (2016), we show how history-dependent labor wedges can implemented with an income-contingent college loan system. Lawson (2016) uses a sufficient statistic approach to characterize optimal uniform tuition subsidies for all college students in a more stylized setting. Our work is also complementary to Abbott, Gallipoli, Meghir, and Violante (2016) and Krueger and Ludwig (2013, 2016) who study education policies computationally in very rich overlapping-generation models. We contribute in this paper by developing a new framework to analyze how education policies should depend on parents' resources and also trade-off merit-based concerns.[7] We are able to characterize optimal financial aid and tax policies theoretically despite allowing for a large amount of heterogeneity and tightly connect our theory directly to the data, by estimating the relevant parameters ourselves.

When characterizing optimal 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 optimal tax papers with both, intensive and extensive margin,

[7]Gelber and Weinzierl (2015) study how tax policies should take into account that the ability of children is linked to parents' resources. Stantcheva (2015) derives education and tax policies in an OLG model with multi-dimensional heterogeneity, characterizing the relationship between education and bequest policies.

as in Saez (2002).[8] More generally, our work is connected to the optimal taxation literature surveyed in Piketty and Saez (2013) and dynamic extensions to characterize more complex policies (Golosov, Tsyvinski, and Werquin 2014). Finally, the paper is also related to many empirical papers, from which we take the evidence to gauge the performance of the estimated model. Those papers are discussed in detail in Section 4.

Our framework makes some simplifying assumptions, which may restrict the generality of our results. First, we abstract from explicitly modeling heterogeneity in college types and majors. Implicitly, sorting of students into different colleges and majors is captured by the estimated differences in returns. Large changes in financial aid policies may change that sorting, and it is conceivable that it would increase the desirability of progressive financial aid if lower income students will select into higher value-added but more expensive institutions. Second, we consider only direct subsidies and do not change student loan policies. However, we show that the issue of loans is slightly orthogonal to the question of progressivity: the policy implications concerning the progressivity do not change if the government in addition provides to opportunity to borrow. Third, we rule out that higher subsidies lead to strategic tuition increases by universities. This may reduce the desirability of public financial aid on average but should only have an impact on the optimal progressivity if it leads to stronger tuition increases for expansion in financial aid for low income children. Further, this should only mute the positive effects in the sense that private colleges respond, where only about 40% of all 4-year college students are enrolled (Snyder and Dillow 2013).

We progress as follows. In Section 2 we develop the model and study optimal policies. Section 3 continues the theoretical part and adds optimal taxation. In Section 4 we describe our calibration and estimation approach and discuss the relationship to previous empirical work. Section 5 presents optimal financial aid policies and Section 6 considers the jointly optimal education and tax policies. In Section 7 we discuss the robustness of our results with respect to college dropout and general equilibrium effects on wages. Section 8 concludes.

# 2 Model and Optimal Financial Aid Policies

The model characterizes optimal financial aid policies for college students. We start by stressing the need-based component of financial aid and derive optimal policies as a function of parental income. Optimal policies will be a function of a set of estimable parameters. In particular,

[8]Kleven, Kreiner, and Saez (2009) consider the extensive margin of secondary earners, Scheuer (2014) considers the occupational choice margin, Saez (2002) and Jacquet, Lehmann, and Van der Linden (2013) consider the labor force participation margin and Lehmann, Simula, and Trannoy (2014) consider migration.

the elasticity of college graduation rates w.r.t. changes in financial aid generosity, the returns for marginal students, and the fraction of inframarginal students will be the key forces driving the most important results. Subsequently, we also allow the government to condition financial aid policies on other observables like academic merit or jointly on the combination of parental income and academic merit.

In the model, individuals start life as high school graduates and decide whether to obtain a college degree. If an individual decides against a college degree, she directly enters the labor market. The decision to enroll into college will depend on a vector of characteristics $X$. For example, potential students may be aware of their returns to college and these returns are likely to be heterogenous. It could also capture geographical origins, endurance or any other aspect that influences the decision to study. In addition to the sources of heterogeneity in $X$, parental income $I$ can determine the college decision. We stress this dimension as an extra parameter because of our strong focus on the need-based element of student financial aid. Parental income $I$ is strongly associated with parental transfers during college. Parental transfers matter for two reasons. First, parental transfers matter because of (potentially binding) borrowing constraints. Second, parental transfers act as a price subsidy because parents make transfers contingent on the educational decision.

The model also incorporates uncertainty about labor market outcomes. We start with a simple two period version of the model with an education period and a labor market period. It is inconsequential for the interpretation of the optimal financial aid formulas, as they also hold if taxable incomes and wages change over the life cycle.

## 2.1 Individual Problem

Individuals graduate from high-school and are characterized by a vector $X \in \chi$ and (permanent) parental income $I \in \mathbb{R}_+$. A certain type $(I, X)$ is also labeled by $j$. They face a binary choice at the beginning of the model: enrolling into college or not. Assume that life after the college entry decision lasts $T$ years, college takes $T_e$ years and individuals' yearly discount factor is $\beta$. Then we can think of $\beta^{C1} = \sum_{t=1}^{T_e} \beta^{t-1}$ and $\beta^{C2} = \sum_{t=T_e+1}^{T} \beta^{t-1}$. If a young individual $j$ enrolls, her expected lifetime utility is:

$$\beta^{C1} U^C(c_j^C; I, X) + \beta^{C2} \int_{\Omega} U^W \left(c_{jw}^W, y_{jw}^W; w, I, X\right) dG^C(w|I, X).$$

$U^C(c_j^C; I, X)$ denotes utility during the college years. It depends on consumption $c_j^C$ during those years, and level of consumption will depend on the realization of $j = (I, X)$. For example higher parental income is strongly associated with higher parental transfers during college. $I$ and $X$ can also have a direct utility effect of attending college; for example empirical studies have found a strong correlation between parents' and children's educational attainment, conditional on parental income. This would be captured by the direct effect of $X$.

The wage $w \in \Omega$ is drawn from a conditional distribution function $G^C(w|I,X)$. $X$ can include, for example, a measure of ability, which leads to heterogeneous returns to college. Empirical paper have stressed the importance of complementarity between ability measures and college education, which can be flexibly captured by $G^C(w|I,X)$. Consumption and taxable income during the working life are $c^W_{jw}$ and $y^W_{jw}$. They depend on the wage draw, as well as the type from the previous period. Parental income $I$ may still influence consumption and labor supply during adulthood, for example, as it determines the need for student loans during college, which are paid back over the working life. The problem of a college graduate with parental income $I$ and vector $X$ becomes:

$$V^C(I,X;\mathcal{G}(I),T(.)) = \max_{c^C_j, c^W_{jw}, y^W_{jw}} \beta^{C1} U^C(c^C_j; I, X) + \beta^{C2} \int_\Omega U^W\left(c^W_{jw}, y^W_{jw}; w, I, X\right) dG^C(w|I,X)$$

subject to

$$\forall\, w:\; c^W_{jw} = y^W_{jw} - T(y^W_{jw}) - (1+r)L$$

and

$$c^C_j = tr^C(I) + \mathcal{G}(I) - \mathcal{C} + L,$$

and

$$L \leq \bar{L}.$$

where $\beta^{C1}$ and $\beta^{C2}$ 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 $\beta$. Then we can think of $\beta^{C1} = \sum_{t=1}^{T_e} \beta^{t-1}$ and $\beta^{C2} = \sum_{t=T_e+1}^{T} \beta^{t-1}$. $r$ is the interest rate that also captures different period length. For example a zero interest rate would imply $1 + r = \frac{\beta^{C1}}{\beta^{C2}}$ in this model.

$T(.)$ are taxes on earnings. $tr^C(I)$ is the transfer function mapping parental income into transfers received when going to college. Students can take loans $L$ with some interest $r$. Potentially, there may be an exogenous borrowing limit on loans taken out given by $\bar{L}$. The government runs a financial aid program $\mathcal{G}(I)$ which subsidizes college costs based on financial needs. $\mathcal{C}$ represents the tuition cost of attending college.

Expected utility of a high-school graduate entering the labor market directly is:

$$\beta^H \int_\Omega U^H \left(c^H_{jw}, y^H_{jw}; w, I, X\right) dG^H(w|I, X),$$

where $\beta^H = \beta^{C1} + \beta^{C2}$ captures the length of the labor market period of high school graduates. The wage realization is drawn from a different conditional distribution $G^H(w|I,X)$, but is allowed to depend on attributes in $X$, importantly ability should be expected to influence wages also for high-school graduates. We will from now refer to all individuals not attending college as high-school graduates. The problem of a high-school graduate with parental income $I$ and vector $X$ becomes:

$$V^H(I, X; T(.)) = \max_{c^H_{jw}, y^H_{jw}} \beta^H \int_\Omega U^H \left(c^H_{jw}, y^H_{jw}; w, I, X\right) dG^H(w|I, X)$$

subject to

$$\forall\, w: \; c^H_{jw} = y^H_{jw} - T(y^H_{jw}) + tr^H(I).$$

So a high-school graduate solves a static problem under this formulation. Note that we also allow for the possibility that high-school graduates receive financial support from their parents $tr^H(I)$. We observe positive transfers in the data also for working high-school graduates and the majority of these transfers happen at the beginning of the working life.

Finally, each type $(I,X)$ decides to attend college or not, comparing $V^C(I, X; \mathcal{G}(I), T(.))$ and $V^H(I, X; T(.))$. We assume that the value functions are differentiable in policies.

## 2.2 Government Problem and Optimal Policies

We now characterize the optimal level of financial aid function $\mathcal{G}(I)$ for a given tax function. We denote by $F(I)$ the unconditional parental income distribution, by $K(I,X)$ the joint c.d.f. and by $H(X|I)$ the conditional one; the densities are $f(I)$, $k(I,X)$ and $h(X|I)$. The support of $I$ and $X$ are $\mathbb{R}_+$ and $\chi$. The government assigns Pareto weights $\tilde{k}(I,X) = \tilde{f}(I)\tilde{h}(X|I)$ which are normalized to integrate up to one. The objective of the government is:

$$\max_{\mathcal{G}(I)} \quad \int_{\mathbb{R}_+} \int_\chi \max\{V^C(I,X), V^H(I,X)\} \tilde{k}(I,X) dIdX \tag{1}$$

s.t. to the budget constraint:

$$\int_{\mathbb{R}_+}\int_{\mathcal{X}} \beta^{C1}\mathcal{G}(I)\mathbb{1}_{V_j^C\geq V_j^H}k(I,X)dIdX = \int_{\Omega}\int_{\mathbb{R}_+}\int_{\mathcal{X}} \beta^{H}T(y_{jw}^H)\mathbb{1}_{V_j^C< V_j^H}k(I,X)dIdXdG^H(w|I,X)$$
$$+\int_{\Omega}\int_{\mathbb{R}_+}\int_{\mathcal{X}} \beta^{C2}T(y_{jw}^C)\mathbb{1}_{V_j^C\geq V_j^H}k(I,X)dIdXdG^C(w|I,X),$$

where $\mathbb{1}_{V_j^C<V_j^H}$ and $\mathbb{1}_{V_j^C\geq V_j^H}$ are indicator functions capturing the education choice for each type $j=(I,X)$. The budget constraint simply equates government spending on financial aid to tax revenues. We label $\rho$ as the multiplier on the budget constraint and assume the government shares the same discount factor as the agents.

Before we derive optimal education subsidies, we ease the upcoming notation a little bit and define the share of college students at parental income level $I$ as follows:

$$F^C(I) = \int_{\mathcal{X}} \mathbb{1}_{V_j^C\geq V_j^H}h(X|I)dX.$$

The marginal impact on welfare of an increase in financial aid $\mathcal{G}(I)$ is given by:

$$\underbrace{\frac{\partial F^C(I)}{\partial \mathcal{G}(I)}\times \Delta T(I)}_{\text{Fiscal Externality}} \quad - \quad \underbrace{F^C(I)(1-W^C(I))}_{\text{Mechanical Effect}}, \tag{2}$$

where $\Delta T(I)$ is the expected fiscal externality (Hendren 2014) from going to college for an average marginal individual with parental income $I$. Formally it is given by

$$\Delta T(I) = \frac{\int_{\mathcal{X}}\mathbb{1}_{H_j\to C_j}\Delta T_j dXh(X|I)}{\int_{\mathcal{X}}\mathbb{1}_{H_j\to C_j}dXh(X|I)} \tag{3}$$

where $\mathbb{1}_{H_j\to C_j}$ takes the value one if individual $j$ is marginal in her college decision with respect to a small increase in financial aid. By definition we have $\int_{\mathcal{X}}\mathbb{1}_{H_j\to C_j}h(X|I)dX = \frac{\partial F^C(I)}{\partial \mathcal{G}(I)}$. Note that the fraction of students $F_I^C$ is differentiable since we assume that the value functions are differentiable in grants. $\Delta T_j$ is the expected fiscal externality of an individual of type $j$:

$$\Delta T_j = \frac{1}{\beta^{C1}}\int_{\Omega}\left(\beta^{C2}T(y_{jw}^C)g^C(w|I,X) - \beta^H T(y_{jw}^H)g^H(w|I,X)\right)dw - \mathcal{G}(I).$$

The first term in (2) captures the fiscal benefits of more financial aid. The reform will trigger enrollment from a certain set of students from income level $I$, those who were close to the margin on enrolling before the reform. This gives rise to the expected increase in tax payment per student type $j$.

The second term in (2) captures the mechanical aspect of the reform: for all inframarginal students at the parental income level in question, the government has to spend one more Dollar since it is impossible to just target marginal students by the reform. The marginal costs are scaled down by the welfare weights on students

$$W^C(I) = \frac{\tilde{f}(I)}{f(I)} \frac{\int_X \mathbb{1}_{V^C \geq V^H} U_c^C(c_j^C; I, X)\tilde{h}(I|X)dX,}{\rho}$$

where $U_c^C$ is the marginal utility of consumption and $\rho$ is the marginal value of public funds – thus, $W^C(I)$ is the money-metric marginal social welfare weight (Saez and Stantcheva 2016).

Multiplying (2) by $\mathcal{G}(I)$ and setting to zero yields a formula for the optimal level of financial aid at parental income $I$:

$$\mathcal{G}(I) = \frac{\eta(I)\Delta T(I)}{F^C(I)(1 - W^C(I))} \tag{4}$$

where $\eta(I)$ is a local elasticity of college enrollment rates: the percentage point change in the share of students in parental income group $I$ in terms of a percentage change in $\mathcal{G}(I)$. Optimal financial aid is increasing in the effectiveness of increasing college attendance measured by $\eta(I)$; such behavioral responses have been estimated in the literature exploiting financial aid reforms, see the discussion in Section 4.3. This behavioral effect is a policy elasticity as discussed in Hendren (2015). This effect is weighted by the fiscal externality created, i.e. the increase in tax payments. Intuitively, the size of the fiscal externality will depend on the returns to college for marginal students, another parameter which has been estimated in different contexts in prior work. Optimal financial aid is decreasing in the number of inframarginal students, capturing the cost of financial aid, and increasing in the value placed on college students' welfare.

The formula is a sufficient statistic formula, providing intuition for the main trade-offs underlying the design of financial aid. It is valid without taking a stand on the functioning of credit markets for students, the riskiness of education decisions or the exact modeling how parental transfers are income influenced by parental income. Changes in those factors would influence the parameter $\eta(I)$, for example, a tightening in borrowing constraints should increase the sensitivity of enrollment especially for low income students.

Notice that the essence of the main trade-offs are unchanged if taxable incomes change over the life-cycle. This affects the calculation of the term $T_j$ which then reflects the difference in discounted present values of yearly tax payments over the life-cycle. Additionally, if wages change stochastically over the life-cycle, the fiscal externality still reflects differences in expected tax payments for the group of marginal students.

**The Role of Parental Income** How should we expect that the optimal $\mathcal{G}(I)$ varies with parental income $I$? On the one hand one should expect a larger effect of increases in financial aid on attendance decisions for low income kids. The size of the fiscal externality $T_j$ is closely related to the returns for marginal students from a parental income group. Ex-ante it is not clear how this term should vary with $I$. The education literature has stressed the complementarity between early childhood human capital investments (Carneiro and Heckman 2003) and found evidence for higher educational returns for children from households with higher income (Altonji and Dunn 1996). It is plausible that this complementarity is also important for marginal students, which suggests higher returns for higher levels of $I$. On the other hand, papers using instrumental variables to estimate returns for marginal students for different kind of policy changes have found relatively large returns (compared to OLS estimates), which is sometimes attributed to high returns for children from economically disadvantaged backgrounds Oreopoulos and Petronijevic 2013).

The RHS of (4) points towards progressive optimal policies given the well-documented correlation between college attendance and parental income (Chetty, Hendren, Kline, Saez, and Turner 2014): the higher parental income, the larger the share of inframarginal students. Additional welfare weights should be plausibly assumed to be decreasing in parents' resources. Our empirical model which we estimate in the next section will shed light on the quantitative importance and magnitudes of these different forces,

Beside the fully optimal level, we will use our empirical model for a related but different question: to what extent could small reforms to the current US financial aid be self-financing through higher future tax-revenue? We consider this as an interesting complementary question for at least two reasons. First, it may be easier to implement small reforms to the existing current federal financial aid system. Second, it points out if there are potential Pareto improving free-lunch policy reforms on the table which are independent of the underlying welfare function.

Setting $W^C(I)$ to 0 to focus on fiscal magnitudes, we can rewrite (2) as:

$$R(I) = \frac{\frac{\partial F^C(I)}{\partial \mathcal{G}(I)} \Delta T(I)}{F^C(I)} - 1 \tag{5}$$

This expression can be interpreted as the rate of return on one dollar invested in additional college subsidies at income level $I$. 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.

It is plausibly of more practical relevance to consider reforms which increases financial aid up for students whose parents' income is below some level $I^*$. For such a reform the fiscal effect is:

$$R^*(I) = \frac{\int_0^I \frac{\partial F^C(I)}{\partial \mathcal{G}(\tilde{I})} \Delta T(\tilde{I}) d\tilde{I}}{\int_0^I F_I^C(\tilde{I}) f(\tilde{I}) d\tilde{I}} - 1, \qquad (6)$$

which is simply the aggregation of the fiscal externalities divided by the fraction of infra-marginals up to income level $I$.

**Comparison to the theoretical literature** The theoretical result that education should be subsidized because of fiscal externalities or – to put it differently – to counteract the distortions of progressive income taxes on the education margin is not new. It has been worked out by Bovenberg and Jacobs (2005) in a static setting and they have shown that under certain conditions on the human capital production function, educational investment in the second-best follows a first-best rule. The optimal education subsidy for each type is set to offset the labor income distortion, which implies that the marginal subsidy rate for each individual is exactly equal to the marginal income tax rate – education subsidies and marginal tax rates are 'siamese twins'. Bohacek and Kapicka (2008) have obtained similar results in a dynamic deterministic environment. Stantcheva (2016) studies a very general dynamic stochastic environment and elaborates the properties of optimal human capital subsidies in second-best efficient allocations. In particular she studies how results differ between the cases where the wage elasticity w.r.t. ability is increasing or decreasing in human capital.

These settings have in common that education subsidies are tailored to different (histories of) types. By contrast, in our setting with an extensive college margin, many different types that face different trade-offs (in terms of returns and constraints) will receive the same subsidy.[9] This changes trade-offs significantly as an increase in subsidies also affects inframarginal types. This makes theoretical predictions for the level of the optimal subsidy less clear. In particular optimal subsidies are rather expressed in absolute numbers and it can well be that they exceed the direct costs (such as tuition fees) of education. But as we showed, optimality conditions depend on simple sufficient statistics that can be mapped to the empirical literature and quantified in a very transparent way.

[9] Jacobs (2007) analyzes optimal education subsidies along the extensive margin in a static deterministic setting with one-dimensional heterogeneity, general equilibrium effects on wages and linear taxes. He finds that education should be taxed on a net basis to tax the rents of infra-marginal students in contrast to the 'siamese-twins' result of Bovenberg and Jacobs (2005). The concept of a 'net tax' is a bit less obvious in our setting with foregone earnings costs of education, uncertainty and nonlinear taxes and the question about its sign – although very interesting from theoretical point of view – is not a question we study in this paper.

An advantage of the extensive margin education model is that it can easily deal with multi-dimensional heterogeneity – this is similar as in the extensive margin labor supply literature (Choné and Laroque 2011). In our quantitative part of the paper, we can therefore account for heterogeneity in parental income, cognitive ability, preferences, borrowing constraints and idiosyncratic wage risk. When we extent the model to also include optimal income taxation, our approach can still accommodate these dimensions of heterogeneity, because they do not affect the labor supply margin in the absence of income effects (Findeisen and Sachs 2015b).

**On the exogeneity of parental transfers** The reader might wonder how problematic our assumption is that parental transfers do only depend on the decision to go to college and parental income but not on grant policies. It has been argued by Abbott, Gallipoli, Meghir, and Violante (2016) that student financial aid tends to crowd-out parental transfers. Certainly, one can expect crowd-out for students that are infra-marginal w.r.t. to their decision. Taking into account for this crowd-out for inframarginal students would barely change our analysis. Solely the social marginal welfare weight $W^C(I)$ would have to be adjusted such that it is taken into account that not each additional dollar of financial aid effectively reaches the student – this would only be a very marginal change of our analysis.

What would be more of an issue is the crowd-out of parental transfers for marginal students. Thus, assuming exogenous transfers, one might overestimate the increase in enrollment if one does not take into account that parents reduce the transfer. Whether such a form of crowd-out really takes place depends on how one models the interaction between parents and children. Assume, e.g., that parents and children decide in a cooperative manner. Such a decision making would in the first step involve the determination of transfers for both situations (college or not) and then in a second step, the decision whether the child goes to college or not. In such a setting, one would never get the result that: (i) parents and the child decide against college for some given level of financial aid, (ii) would decide in favor of college after an increase in financial aid for parental transfers as in (i), (iii) would decide against college after adjusting the parental transfers conditional on college to respect the increase in financial aid. If (ii) holds, then they would certainly be better of than before the grant increase. Thus, it cannot be optimal to adjust the college transfers such that the child does not go to college. This would leave the joint decision makers equally well of as before the grant increase.

Our point here is not to argue that crowd-out is no issue at all – a different modelling of the parental transfer decision has different implications. In fact an extension of our framework to endogenize parental transfers would be relatively straightforward. But let us highlight the sufficient statistic nature of our results at this point. Later in our quantitative section we calibrate the model to match the increase in college graduation due to an increase in financial aid. The quasi-experimental elasticities that we target there are changes in equilibrium outcomes (that take into account potential crowd-out in parental transfers). Thus, if we were to extend the model by endogenizing parental transfers, we would nevertheless target these numbers in the calibration and would obtain (by construction) the same local elasticities. What might change, however, are the the college enrollment elasticities for policies that are further away from current policies. Given that Abbott, Gallipoli, Meghir, and Violante (2016) find that crowd-out is rather an issue for students with better financial background, we do conjecture that such an extension would reinforce our result on the optimal progressivity of financial aid.

## 2.3 Merit-Based Policies

Our approach is more general and can be extended to condition financial aid policies on other observables like academic merit or jointly on the combination of parental income and academic merit. In fact in our empirical application we will allow the government to also target financial aid policies on a signal of academic ability. Suppose the government can observe such a signal of academic ability like the SAT score. We take that factor out of the vector $X$ and label it $\theta$. For notational simplicity, we will still call the vector without $\theta$ $X$; in this case $X$ includes all factors influencing the college decision except for parental income and the measure of academic ability. Suppose we are interested in deriving the optimal policy schedule which conditions on need- and merit-based components jointly. Formally, the government maximizes over $\mathcal{G}(I, \theta)$. The derivation of the optimal financial aid policy schedule is analogous to the derivation of $\mathcal{G}(I)$ and yields:

$$\mathcal{G}(I, \theta) = \frac{\eta(I, \theta)\Delta T(I, \theta)}{F^C(I, \theta)(1 - W^C(I, \theta))}, \tag{7}$$

where all terms are evaluated at a parental income-ability pair $(I, \theta)$.

How should we expect optimal financial aid expect to vary with academic ability, holding parental income fixed? At first glance, one may expect that the optimal grant $\mathcal{G}(I, \theta)$ is increasing in $\theta$ as the returns to college education should increase in $\theta$, which boosts the fiscal externality. By conditioning on ability directly, the government can implicitly guarantee that marginal students have a certain minimum expected return to college attendance, circumventing some of the potential problems of a pure need-based system. Working against this is that higher ability students are likely more inframarginal in their decision: i.e. they opt for college in any financial aid system. Our empirical model will shed light on this first question, which has no clear theoretical answer.

# 3 Optimal Taxation

Our previous analysis hinted at the importance of income taxation for the design of optimal financial aid policies. We now extend the model to allow the government to also chose income taxation optimally. We consider this is an important extension for three reasons. First, as the last section has shown, higher and more progressive taxes are a complement to financial aid. The average level and also the progressivity of financial aid are hence closely related to the design of taxes.[10] Second, financial aid conditioning on parental resources is partly a redistribution device, captured by the welfare weights in formula (4). When we allow the government to design the optimal non-linear tax system, we can analyze how much of the progressivity of financial aid is driven by the desire to ex-ante redistribute. Finally, we can theoretically and empirically analyze how taxes themselves may distort education decisions, a channel analyzed in a prominent paper by Trostel (1993) .[11]

We build on the large literature following Mirrlees (1971) and the modern literature originating with Diamond (1998) and Saez (2001) expressing optimal tax schedules in terms of observables (see Piketty and Saez (2013) for a review). Our model can stay very general in terms of the underlying heterogeneity, while still preserving tractability.

The planner's problem is the same as in (1) with the difference that the planner also optimally chooses the income tax schedule $T(\cdot)$. Notice that the formula for optimal financial aid policies is unaltered. We allow the tax function $T(\cdot)$ to be arbitrarily nonlinear in the spirit of Mirrlees (1971). We restrict the tax function to be only a function of income and to be independent of the education decision. This tax problem can either be tackled with a variational or tax perturbation approach (Piketty 1997, Saez 2001, Golosov, Tsyvinski, and Werquin 2014, Jacquet and Lehmann 2016) or with a restricted mechanism design approach for nonlinear history-independent income taxes that we explore in Findeisen and Sachs (2015b).

**Assumption** Preferences $U^H\left(c^H_{jw}, y^H_{jw}; w, I, X\right)$ and $U^W\left(c^W_{jw}, y^W_{jw}; w, I, X\right)$ imply no income effects on labour supply.

[10] Bovenberg and Jacobs (2005) was the first paper to emphasize this complementarity. They study a case with a continuous education choice in which the optimal education subsidy rate is equal to the tax rate.

[11] See Abramitzky and Lavy (2014) for recent quasi-experimental evidence on the negative effect of redistributive taxation on education investment.

As we show in Appendix A.2, the optimal marginal tax rate can be expressed as:

$$\frac{T'(y(w^*))}{1-T'(y(w^*))} = \frac{1}{\varepsilon_{y(w^*),1-T'}} \times \left( \mathrm{Haz}(y(w^*))\,(1-\mathcal{W}(y(w^*))) + \frac{\int_{\mathbb{R}_+} \xi(I,y)\Delta T(I,y)dF(I)}{h(y(w^*))y(w^*)} \right) \quad (8)$$

where

$$\mathrm{Haz}(y(w^*)) = \frac{\int_{y(w^*)}^{\infty} h(y)dy}{h(y(w^*))y(w^*)}$$

and

$$\begin{aligned} h(y(w^*)) =& \beta^{C2} \int_{\mathbb{R}_+} \int_{\mathcal{X}} \mathbb{1}_{V^C \geq V^H} k(I,X) g^C(w^*|I,X) dIdX \\ &+ \beta^H \int_{\mathbb{R}_+} \int_{\mathcal{X}} \mathbb{1}_{V^H \leq V^H} k(I,X) g^C(w^*|I,X) dIdX. \end{aligned}$$

Note that $\mathrm{Haz}(y(w^*))$ and $h(y(w^*))$ are basically the Hazard ratio (Saez 2001) and the density of income, only adjusted by period length. $\mathcal{W}(w)$ is a money metric average social welfare weights of individuals with wage $w$. $\varepsilon_{y(w^*),1-T'}$ is the local labor supply elasticity along a nonlinear tax schedule (Jacquet and Lehmann 2016). To capture the college responses to taxes, we define

$$\xi(I,y) = \frac{1}{f(I)} \frac{\partial F_I^C}{\partial T(y)},$$

which is the semi-elasticity of enrollment with respect to the absolute tax at income $y$.

$\Delta T(I,y)$ is the average fiscal externality of those students with parental income $I$ that are marginal w.r.t. a small increase in $T(y)$. It is different to (3), where the average was taken over all students that are marginal wr.t. a small increase in financial aid.

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

$$\int_{\mathbb{R}_+} \xi(I,y)\Delta T(I)dF(I)$$

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.[12] In these papers, the extensive margin is an unambiguous force towards

[12]Further papers are Scheuer (2014) where the extensive margin captures the decision to become an entrepreneur and Kleven, Kreiner, and Saez (2009) who consider the extensive margin of secondary earner to study the optimal taxation of couples.

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 $\Delta T(I, y)$ in general. Second, we do not know whether higher taxes for individuals with $w > w^*$ indeed lead to lower college enrollment because of possibly counteracting income and substitution effects. Whereas higher taxes unambiguously decrease the return to college, an income effect on college enrollment might work in the opposite direction. Further, higher taxes decrease the opportunity costs from going to college in the form of foregeone earnings. In an earlier version of this paper, we distinguish these effects more formally. (Findeisen and Sachs 2015a, p.12) Whether and to what extent the endogeneity of college enrollment leads to lower optimal marginal tax rates is thus a quantitative question.

# 4 Estimation and Calibration

We first explain how we concretely specify the model in Section 4.1. In Section 4.2 we explain how we quantify the model using micro data and information on current policies. In Section 4.3 we show in detail that the quantitative model performs very well in replicating patterns in the data and quasi-experimental evidence on returns to college and the elasticity of college education with respect to financial aid.

## 4.1 Empirical Model Specification

We now specify the concrete set-up for the empirical model. Concerning the underlying heterogeneity, we specify the vector $X$ as $(\theta, \kappa)$, where $\theta$ is ability and $\kappa$ are psychic costs. In the estimation, $\kappa$ will also be allowed to depend on parental education. We assume that ability directly influences the wage distribution, i.e. we specify the wage distributions as $G^C(w|\theta)$ and $G^H(w|\theta)$. We assume these functions to be independent of parental income because we did not find a strong significant effect of parental income. In the modeling of psychic costs, we closely follow the structural education literature; see, among others, Cunha, Heckman, and Navarro (2005), Heckman, Lochner, and Todd (2006), Cunha, Karahan, and Soares (2011), Navarro (2011) and Johnson (2013). Psychic costs an be interpreted as a one-dimensional aggregate that captures factors that influence the decision to go to college beyond the budget constraint. We borrow the notion psychic costs from the empirical literature. They enter the model in a vary simple way: $\kappa$ is just subtracted from lifetime utility if an individual goes to college. The value functions in case of college attendance is

$$V^C(I,\theta,\kappa;\mathcal{G}(.),T(.)) = \max_{c_j^C, c_{jw}^W, y_w^W} \beta^{C1} U^C(c_j^C) + \beta^{C2} \int_\Omega U^W\left(c_{jw}^W, y_w^W; w\right) dG^C(w|\theta) - \kappa$$

subject to

$$\forall\, w:\; c_{jw}^W = y_{jw}^W - T(y_w^W) - (1+r)L$$

$$c_j^C = tr^C(I) + \mathcal{G}(.) - C + L,$$

$$L \leq \bar{L},$$

where $j$ is a realization the triple $(I,\theta,\kappa)$. So we make the standard assumption that preferences over consumption and work are homogenous. Consumption in college differs because of heterogeneity in parental transfers, financial aid receipt, and borrowing. We assume that agents are borrowing constrained and can only borrow up to $\bar{L}$ but show that our implications are not altered if agents can freely borrow. For high-school graduates, we have:

$$V^H(I,\theta;T(.)) = \max_{c_w^H, y_w^H} \beta^H \int_\Omega U^H\left(c_w^H, y_w^H; w\right) dG^H(w|\theta)$$

subject to

$$\forall\, w:\; c_w^H = y_w^H - T(y_w^H) + tr^H(I).$$

Note that the notation implies the fact that taxable income $y$ only depends on $w$ because of Assumption 1. We assume the utility function to be of the following form for both high-school and college graduates.

$$\frac{\left(C - \frac{\left(\frac{y}{w}\right)^{1+\varepsilon}}{1+\varepsilon}\right)^{1-\gamma}}{1-\gamma}.$$

During college, $l$ is set to 0. We choose $\varepsilon = 2$, which implies a compensated labor supply elasticity of .5[13] Note that the value of the labor supply elasticity does not influence our results for given taxes because we calibrate wages from elasticities and income as in Saez (2001). We are more explicit about that in Section 4.2.2. The value of the curvature parameter $\gamma$ matters for the elasticity of the college education decision. We set $\gamma = 1.85$ as this implies an elasticity in the mid range of estimates from the empirical literature. We comment on that more in Section 4.3.

[13]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.

An important simplifying assumption we make is that we abstract from the direct modeling of labor supply behavior over the life-cycle, as we are mostly interested in getting the net-present value of the fiscal externalities over the life-cycle right. This is achieved by using annuity values of the average discounted sums of income, as we describe below. Such simplifications are also commonly made in other calculations, calculating the lifetime present value effects of policies on earnings in the literature, for example, Kline and Walters (2016) and Chetty, Friedman and Rockoff (2014). Under our assumptions of no income effects, this simplification should be of no consequence for the quantitative results. If there are wealth effects on labor supply, a student debt channel could potentially affect our optimal policy results. Suppose changes in financial aid change the borrowing behavior and the amount of student debt carried over. This would differentially affect the labor supply behavior of low-income children with higher debt relative to high-income children with low debt. This effects would probably be mostly present at the beginning of the working life.

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. 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).[14] We set the risk free interest rate to 3%, i.e. $R = 1.03$ and assume that individuals' discount factor is $\beta = \frac{1}{R}$.

## 4.2 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 of ability. Cunha, Karahan, and Soares (2011), e.g., show that it is the most precise signal of innate ability among comparable scores in other data sets.

To quantify the joint distribution of parental income and ability, we take the cross sectional joint distribution in the NLSY97. 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

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

Table 1: Quantification of the Model

| Object | Description | Procedure/Target |
|---|---|---|
| $F(I)$ | Marginal distribution of of parental income | Directly taken from NSLY97 |
| $(\theta, I)$ | Joint and conditional distribution of innate abilities | Directly taken from NSLY97 |
| $w$ | Individual wage | Calibration from income as in Saez (2001) |
| $G^H(w\|\theta)$ | Conditional Wage Distribution High-school | Estimated from regressions |
| $G^C(w\|\theta)$ | Conditional Wage Distribution College | Estimated from regressions |
| $tr^H(I)$ | Conditional Transfer Distribution High-school | Estimated from regressions |
| $tr^C(I)$ | Conditional Transfer Distribution College | Estimated from regressions |
| $K(\theta, I, \kappa)$ | Joint distributions with psychic costs | Maximum Likelihood |
| | Utility Function: $\frac{\left(C-\frac{l^{1+\varepsilon}}{1+\varepsilon}\right)^{1-\gamma}}{1-\gamma}$ | |
| $\varepsilon$=0.5 | Labor Supply Elasticity | Chetty, Guren, Manoli, and Weber (2011) |
| $\gamma$ =1.85 | Curvature of Utility | Enrollment Elasticitites |
| | Current Policies | |
| $\bar{s}$ | Stafford Loan Maximum | Value in year 2000 |
| $T(y)$ | Current Tax Function | Gouveia-Strauss (Guner et al. 2013) |
| $\mathcal{G}(\theta, I)$ | Need- and Merit Based Grants | Estimated from regressions |

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. 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 (2016). The underlying assumption is that the relation between AFQT and wages has not changed over that time period. 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 take current policies as described below. We then proceed as follows:

1. We estimate $G^H(w|\theta)$, $G^C(w|\theta)$ in 4.2.2.

2. Transfer function $tr^C(I)$, $tr^H(I)$ and grant receipt are estimated $\mathcal{G}(\theta, I)$ by regressions. So we estimate empirically the need-based and merit-based component of current financial aid. For brevity, details of our procedure for transfers and grants are relegated to Appendix A.3. Economically, the most important results for parental transfers is the strong dependence on education choice by the child. This contingency of parental transfers acts

as a price subsidy for college. On top, we recover the well-know positive correlation between parental income and transfers. For grants, we find a strong negative effects of parental transfers on financial aid receipt at the extensive and intensive margin. Additionally, we can capture merit based grants by the conditional correlation of AFQT scores with grant receipt.

3. Based on that, we calculate $V^C(I, \theta, \kappa; \mathcal{G}(.), T(.))$ and $V^H(I, \theta; T(.))$ for each individual and estimate the distribution of psychic costs with maximum likelihood in Section 4.2.3.

### 4.2.1 Current Policies

To capture current tax policies, we use the approximation of Guner, Kaygusuz, and Ventura (2014), which has been shown to work well in replicating the US tax code. More details are contained in Appendix A.3.1. 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.3.2. The average values are $7,434 for annual tuition and $4,157 for the annual subsidy (public appropriations) 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 A.3.5 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 is targeted to the respective quasi-experimental evidence. Further, we show that our main result about the progressivity of optimal financial aid prevails if we allow for free borrowing in Section 5.3.

### 4.2.2 Estimation of Wage Functions

In our model, $y$ refers to an average income over the lifetime as we only have one labor market period. Therefore, we took annuitized income as the data counterpart. Our approach to estimate the relationship between innate ability, education and labor market outcomes relates to Abbott, Gallipoli, Meghir, and Violante (2016) and Johnson (2013). We run regressions of log annuitized income on AFQT for both education levels. This gives us conditional log-normal distributions of labor income. See Appendix A.3.3 for details

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.[15] Figure 11(a) in Appendix A.1 shows the 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 2003). The red bold line in Figure 11(b) in Appendix A.1 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 A.3.5. The net impact on public funds is given by the blue dashed-dotted line.

The last step consists of calibrating the respective skill/wage distribution from the income distributions by exploiting the first-order condition of individuals as pioneered by Saez (2001). This highlights that our results for an exogenous tax function are independent of the labor supply elasticity. The wages are always calibrated such that they produce the income distribution that we estimated. If we change the value of the elasticity, the wages adjust accordingly. For our results on optimal taxes, the labor supply elasticity matters of course – the higher it is the lower are optimal taxes. However, it does not have significant consequences for the optimal progressivity of financial aid as we find in unreported simulation exercises.

### 4.2.3 Estimation of Psychic Costs

Based on the estimated reduced form relationships, we can calculate the two value functions for each individual in the data. In line with the empirical literature, we assume that the decision to go to college is also influenced by heterogeneity in preferences for college. We assume that

[15] 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 did not find any systematic relationship between the Pareto parameter and either $\theta$ or education in the NLSY.

these psychic costs are determined by parental education and by innate ability – see Cunha, Heckman, and Navarro (2005), among others.[16] To achieve identification we impose a normality assumption on the distribution of preferences. The model is estimated with maximum-likelihood and details of the procedure are found in Appendix A.3.6.

## 4.3 Model Performance and Relation to Empirical Evidence

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.

**Graduation Shares.** Figure 10 in Appendix A.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. 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 30.85% 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.

**Responsiveness of Graduation to Grant Increases.** Many paper have analyzed 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

[16] 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.

of Snyder and Dillow (2013)). Of those 66%, only slightly more than half should be expected to graduate with a bachelors degree. 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 (Shapiro et al. 2012, Figure 6). 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. We chose the parameter $\gamma = 1.85$ of the utility function such that we are exactly in the middle of this range at 1.6.

A more recent study by Castleman and Long (2015) 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 after 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 10(a). 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 (2015) 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 0.08 percentage points. In line with Hilger (2015), our model predicts a very moderate effect of parental income, smaller but in line with Hilger (2015).

**The College Wage Premium and Marginal Returns.** The college-earnings premium in our model is 99%, 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 not gone to college, we find that the returns to college are 62.9%. 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).[17]

The more important number for our analysis is the return to college for *marginal students*. We find it to be slightly lower at 58.62%, which implies a return to one year of schooling of 11.53%. This reflects that marginal students are of lower ability on average than inframarginal students and also is in line with Oreopoulos and Petronijevic (2013). A clean way to infer returns for marginal students is found in 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.[18] 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 51.73% for students with a GPA in the neighborhood of 3, which implies a Mincer return of 10.42% for one year of schooling – which comes very close to the 9.9% from Zimmerman (2014).

Finally, 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 the case for an increase in college subsidies.

**The Role of Borrowing Constraints.** To assess the importance of borrowing constraints, we completely remove them to ask by how much graduation increases. In this experiment enrollment increases by 3.94 percentage points from 30.85% to 34.79%. This value is in the realm of values the literature has found, see, e.g., Johnson (2013) and Navarro (2011). As Figure 12(a) in Appendix A.1 reveals, the removal of borrowing constraints has larger effects

[17]The calculation is as follows. In a Mincer regression, the log of earnings is regressed on years of schooling. The difference in $log(1.64y)$ and $log(y)$ is equal to $log(1.64)$. Dividing by four years of schooling (for a bachelor degree) yields 12.20% per year of schooling.

[18]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.

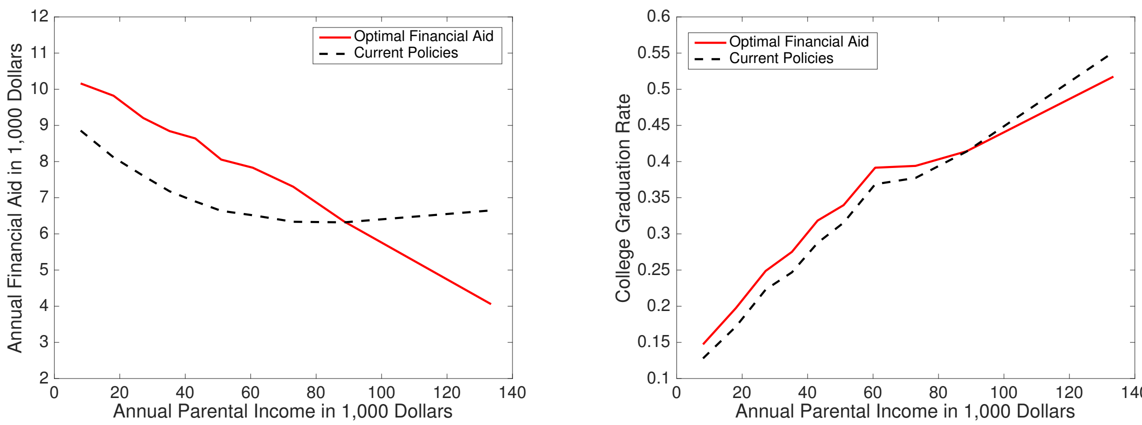
for low income children. Figure 12(b) in Appendix A.1 illustrates the importance of borrowing constraints for individuals with different innate abilities. Naturally, individuals with high ability have the strongest need for more borrowing because of high expected future earnings.

# 5 Results: Optimal Financial Aid

In this section we quantitatively present our main result of progressive financial aid policies. After presenting the benchmark in Section 5.1, we show that results are robust to the welfare function and also hold if the government only wants to maximize tax revenue in Section 5.2. One might think that results are driven by borrowing constraints. As we show in Section 5.3, even if a perfect credit market could be provided, the optimal financial aid schedule is strongly progressive. In Section 5.4, we also chose the need-based element optimal and find that this does not at all alter our result. We show a larger degree of progressivity can be implemented in a Pareto improving way in Section 5.5.

## 5.1 Optimal (Need-Based) Financial Aid

For our first policy experiment, we ask which levels of financial aid for different parental income levels maximize welfare and thus solve (4). For this experiment, we do not change taxes or any other policy instrument but instead only vary the targeting of financial aid. At this stage, we leave the merit-based element of current financial aid policies unchanged, i.e. we do not change the gradient of financial aid in merit. In Section 5.4, we show that our main result also extends to the case where the merit-based elements are chosen optimally.




(a) Financial Aid (b) Graduation Rates

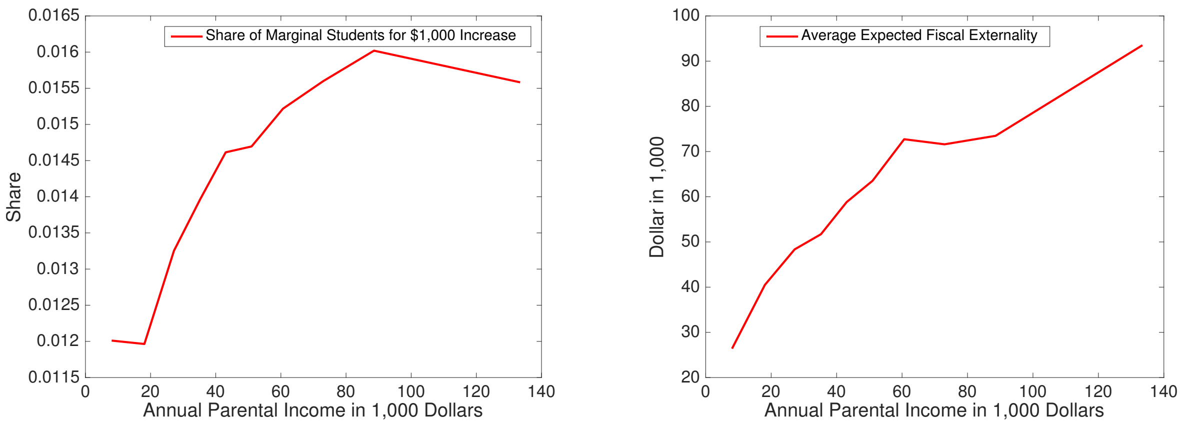
Figure 1: Optimal versus Current Financial Aid

Figure 1(a) illustrates our main result for the benchmark case. Optimal financial aid is strictly decreasing in parental income. Compared to current policies, financial aid is higher for students with parental income below $90,000. This is partly financed by a reduction of financial aid for richer students and partly by the fact that the increase is more than self-financing for the poorest students as we further elaborate in Section 5.5. This change in financial aid policies is mirrored in the change of college graduation as shown in Figure 1(b). The total graduation rate increases by 1.6 percentage points to 32.44%. This number highlights the efficient character of this reform.

**Why Are Optimal Policies So Progressive? A Decomposition.** We now illustrate what drives the progressivity result. From the optimality condition

$$\frac{\partial F^C(I)}{\partial \mathcal{G}(I)} \times \Delta T(I) - F^C(I) \times (1 - W^C(I)) = 0$$

we plot each of the components evaluated the optimal system. Figure 2(a) plots the share of marginal students $\frac{\partial F^C(I)}{\partial \mathcal{G}(I)}$ against parental income in the optimal system. It actually shows

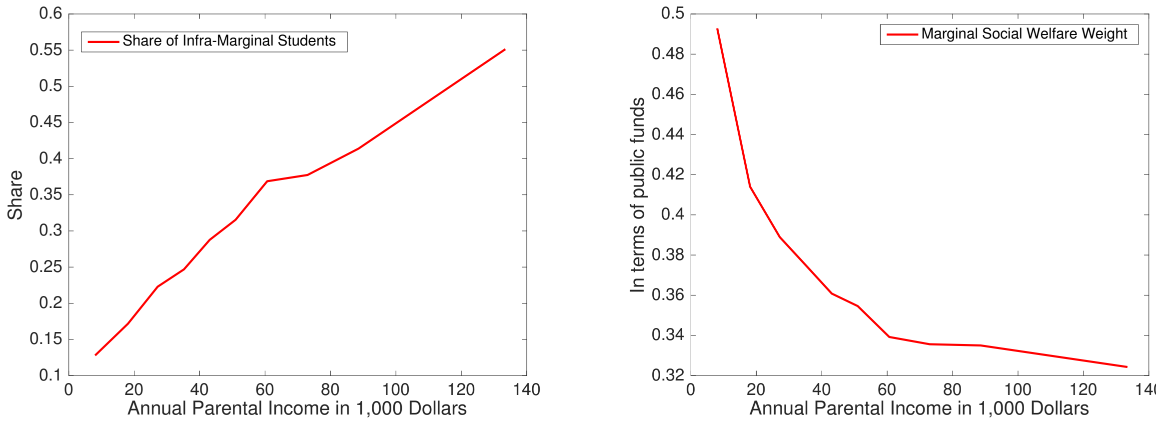



(a) Marginal Students

(b) Fiscal Externality

Figure 2: Marginal Students and Fiscal Externality in Optimal System

an increasing share of marginal students but the relative differences are small as the share increases from 1.2% to around 1.6%. This works against our progressivity result. Figure 2(b) shows the implied average fiscal externality at the optimal system. It increases by a factor around 3 from $30,000 to $100,000. This implies that also the shape of $\Delta T(I)$ works against the progessivity result because marginal students from higher income households have higher returns. Figure 3(a) plots the share of inframarginal students, showing that even in the optimal system there is a strong parental income gradient, as the share increase from around 12% to



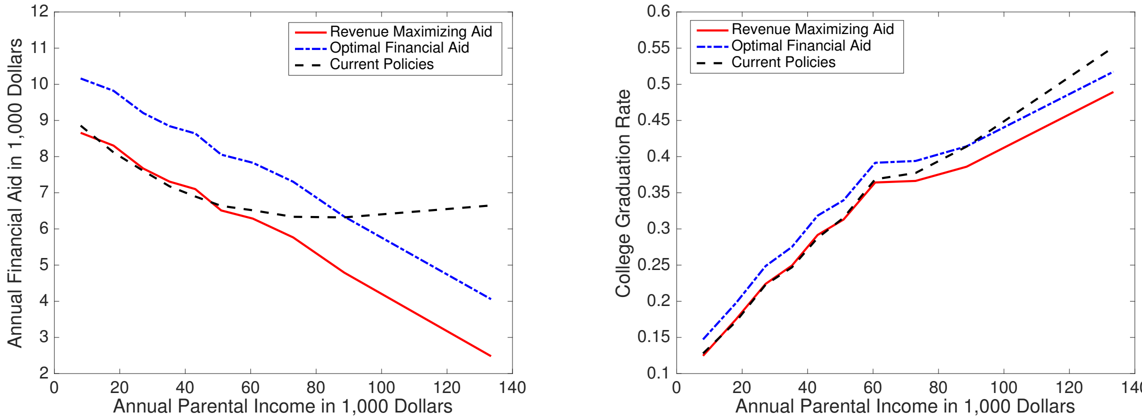

(a) Inframarginal Students (b) Marginal Welfare Weights

Figure 3: Inframarginal Students and Welfare Weights in Optimal System

around 55% implying a factor of around 4.5. Finally, Figure 3(b) shows the implied marginal welfare weights at the optimum. They imply that $1-W^C(I)$ which is the relevant term for the formula increases from around 0.5 to around 0.7 at the top, so by a factor of around 1.4. Taken together, the decomposition yields that the share of inframarginal students is key to explain the progressivity result. Although marginal students from higher incomes have higher returns to college, working against progressive aid policies, this is overturned by the fact that a college attendance is still highly correlated with parental resources. Put differently, even though a progressive system subsidizes low income children much more, high income children are still more likely to attend college.

## 5.2 Tax-Revenue Maximizing Financial Aid

One might be suspicious that the progressivity is driven by a desire for redistribution from rich to poor students. If this were the case, the question would naturally arise whether the financial aid system is the best means of doing so. However, we now show that the result even holds in the absence of redistributive purposes. We ask the following question: how should a government that is only interested in maximizing the budget set financial aid policies? Figure 4(a) provides the answer: revenue maximizing financial aid in this case is very progressive as well. Whereas the overall level is naturally lower if the consumption utility of students is not valued, the declining pattern is basically unaffected. For lower parental income levels, revenue maximizing aid is even above the current one which implies that an increase must be more than self-financing. We study this in more detail in Section 5.5. The implied graduation patterns are illustrated in Figure 4(b).




(a) Financial Aid (b) Graduation Rates

Figure 4: Tax Revenue Maximizing Financial Aid Policies

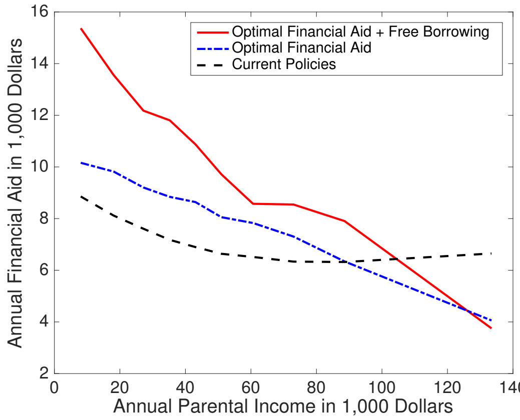
## 5.3 The Role of Borrowing Constraints

We have shown that the optimal progressivity is not primarily driven by redistributive tastes but rather by efficiency considerations. Given that our analysis assumes that students cannot borrow more than the Stafford Loan limit, the question arises these efficiency considerations are driven by borrowing limits that should be particularly binding for low parental income children.

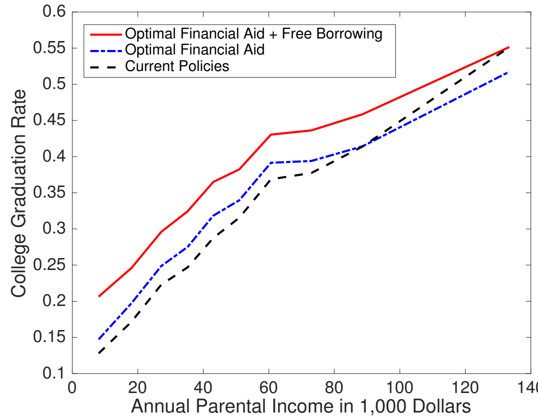
To elaborate upon this question, we ask how normative prescriptions for financial aid policies change if students can suddenly borrow as much as they want. As illustrated in Figure 5(a), optimal financial aid policies become even more progressive in this case. The abolishment of borrowing constraints implies a boost in college education which implies a large increase in tax revenue that can now be used to increase financial aid. The increase is mainly targeted at the low parental income children. First because of their higher welfare weight. Second because also in the absence of borrowing constraint, the general force survives that subsidizing low parental income children is relatively cost-effective because of the much lower share of inframarginal students as can be seend in Figure 5(b).

## 5.4 Merit Based Financial Aid

Up to now, we have assumed that the merit-based element of financial aid policies stays unaffected. We now allow the government to optimally choose the gradient in merit and parental income. Figure 6(a) shows that – if optimally targeted also in terms of merit – financial aid policies can be more generous. The progressive nature however is even slightly reinforced.

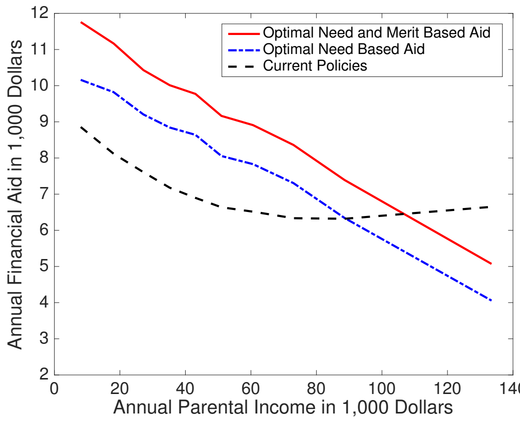



(a) Financial Aid

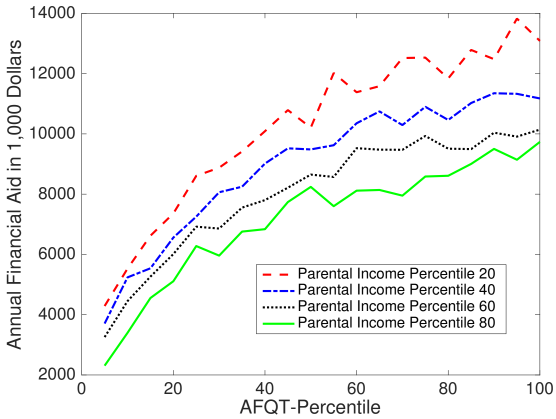



(b) Graduation Rates

Figure 5: Financial Aid and Graduation with Free Borrowing




(a) Financial Aid




(b) Illustration of Optimal Merit Based Element

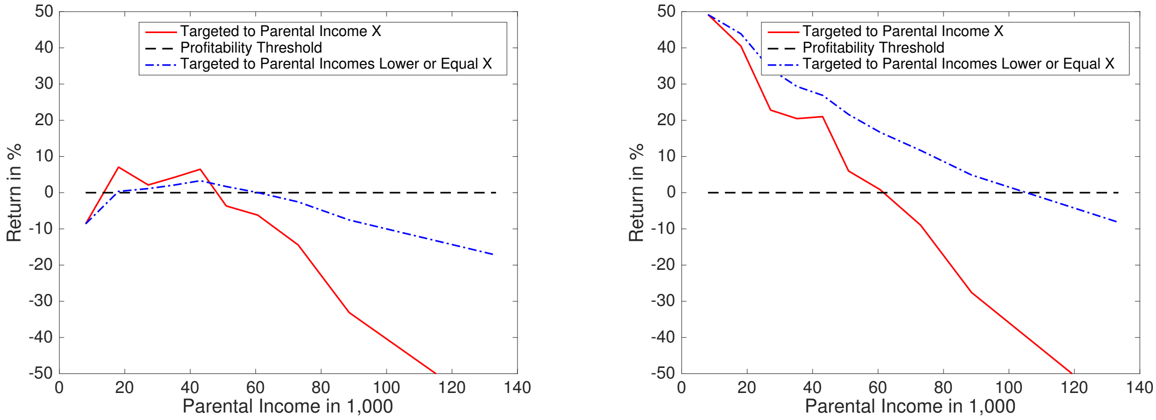
Figure 6: Optimal Need and Merit Based Financial Aid

Figure 6(b) shows how optimal financial aid is increasing is increasing in AFQT. Interestingly, the slope is almost independent of parental income.

## 5.5 Pareto Improving Reforms

As anticipated in Section 5.2, an increase in financial aid can be self-financing if properly targeted. The red bold line in Figure 7(a) illustrates the fiscal return as defined in (5), i.e. if financial aid is increased for a particular income level. Returns are positive between parental income \$18,000 and \$43,000 reflecting roughly the 15% and 45% percentile. This is a striking result: increasing subsidies for this group is a free lunch. An alternative would be to consider reforms where financial aid is increased for students below a certain parental income level which refers to equation (6) in our theoretical part. This case is illustrated by the blue dashed-dotted line in Figure 7(a). An increase in financial aid targeted to children with parental income below

$60,000 is slightly above the margin of being self-financing. Figure 7(b) illustrates the same, however for the case where subsidies are only increased for those AFQT scores above the 50th percentile. Here policy implications become more stark. An increase in subsidies targeted to the poorest students can have a huge fiscal returns of up to 50% as defined in equation (6). Thus, for each marginal dollar invested in grants the government obtains $1.50 in discounted future tax revenue.




(a) All Ability Levels

(b) Targeted to Above Median Ability

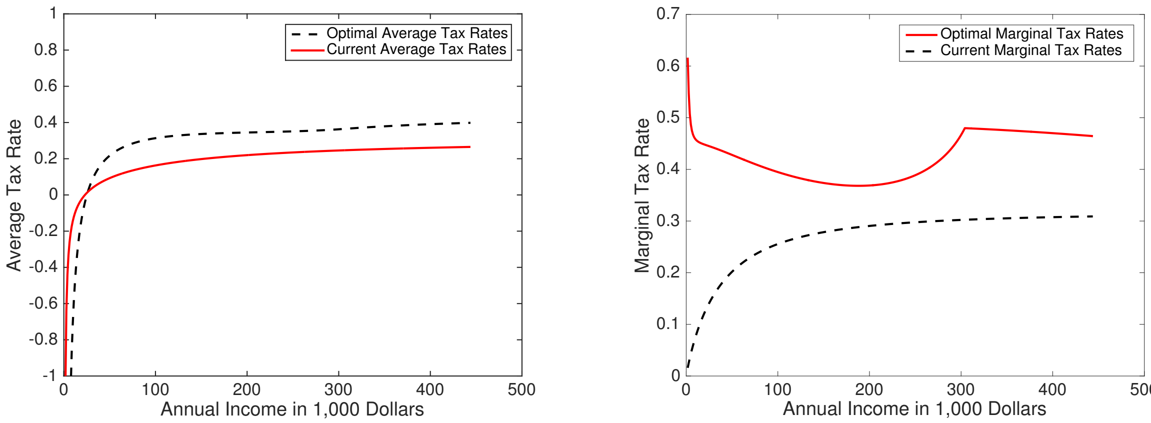
Figure 7: Fiscal Returns on Increase in Financial Aid

# 6 Results: Jointly Optimal Financial Aid and Income Taxation

The previous section has shown that optimal financial aid policies are very progressive. In particular, we have emphasized the efficiency role of progressive policies. Nevertheless, one might wonder how robust this result is with respect to the tax system. Given that the optimal Utilitarian tax schedule is likely to be more progressive than the current tax schedule, how do our results for optimal financial aid change if the tax system is chosen optimally? In particular, is there a trade-off between ex-ante redistribution (through progressive financial aid policies) and ex-post redistribution (through progressive income taxation)? In Section 3, we have shown how to theoretically tackle the issue in the spirit of Mirrlees. Thus, we allow the tax function to be arbitrarily nonlinear. We assume that agents are borrowing constraint and the government only (besides the tax schedule) maximizes the need-based element of the financial aid schedule. Results are barely changed if borrowing constraints are relaxed and/or the merit-based element is chosen optimally as well.

Figure 8(a) displays optimal average tax rates in the optimal as well as the current US system. Average tax rates are higher for most part of the income distribution. As Figure 8(b) shows this is driven by higher marginal tax rate throughout but especially at the bottom of the distribution, a familiar result from the literature (Diamond and Saez 2011). In unreported results we find that the direct of taxes on enrollment decisions, which we discussed in Section 3, is very small. In particular, it does not overturn the optimal U-shaped pattern of optimal tax rates nor does it influence the optimal top tax rate which is still mainly determined by the interaction of the labor supply elasticity and the Pareto parameter of the income distribution (Saez 2001).

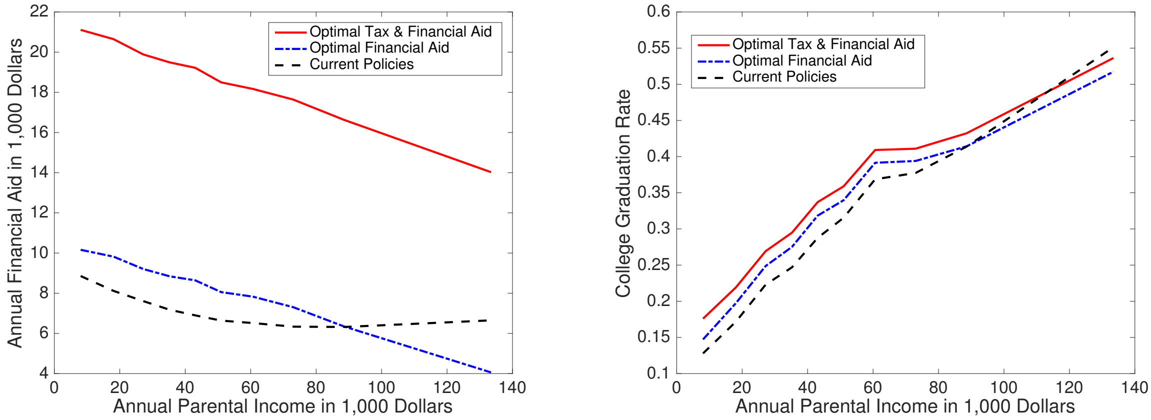
Figure 9(a) illustrates optimal financial aid in the presence of the optimal tax schedule. First notice that financial aid is significantly higher on average compared to case with the current US tax code. Higher income tax rates increase the fiscal externality, which increases the optimal level of the college subsidy (i.e. financial aid). Second, strinkingly the progressivity of optimal financial aid policies is preserved. Progressive taxation does not change the desirability of progressive financial aid policies. A decomposition exercise as in Section 5 show that this is again driven by the increasing share of inframarginal students along the parental income distribution. In other words, many more children from higher income households are inframarginal in their college decision in *any* financial aid or tax system.




(a) Average Tax Rates

(b) Marginal Tax Rates

Figure 8: Optimal versus Current: Average and Marginal Tax Rates




(a) Financial Aid

(b) Graduation Rates

Figure 9: Financial Aid and Graduation with Optimal Redistribution

# 7 Further Aspects

In this section we argue that the result about the progressive nature of optimal student financial aid are unlikely to change if college dropout and general equilibrium effects on wages are taken into account.

## 7.1 Dropout

In our analysis we assumed that anybody who goes to college indeed graduates. Shapiro et al. (2012, Table 6) document 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. So one might wonder to what extent our results are robust to the incorporation of dropout. If one things about the optimality condition (4), what is changed? (i) The marginal costs of the reform are increased because the increase in subsidies now must also be paid for students that are inframarginal but do not graduate. If, for example, 38% are dropouts and the stay on average in college for two years, the marginal cost term – abstracting from discounting – is increased by $\frac{0.38}{0.62}\frac{2}{4.5}$ by 50%. (ii) An increase in college subsidies does not only imply marginal students that graduate but also marginal students that dropout. Note that for our quantitative part we were not making the mistake of assuming that every additionally enrolled student graduates. Instead we were only taking into account the share of those that actually graduate, see also our discussion in Section 4.3. Taking into account that higher subsidies in addition induce marginal students that dropout might make a an increase of grants more or less desirable depending on whether the college dropouts contribute more to public funds over their lifecycle than they would have in the absence of any college education. According to

Lee, Lee, and Shin (2014) the earnings premium for 'some college' was between 25% and 40% between 1980 and 2005. In an earlier version of that paper (Findeisen and Sachs 2015a) we extended our marginal reform approach to incorporate these two aspects of dropout. We found that overall the desirability to increase grants is muted by dropout but did not find that it significantly changed the result that increasing grants for students with low parental income yields higher fiscal returns than for the average. However, there is a third effect that we did not take into account and which should reinforce the progressive nature of optimal financial aid. College grants increase persistence, in particular for students with weak parental background (Angrist, Autor, Hudson, and Pallais 2015, Bettinger 2004, Castleman and Long 2015). This effect would reinforce our normative implications about the progressivity of financial aid.

## 7.2 General Equilibrium Effects on Wages

Our analysis 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 generally 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. But there is a counteracting force: the increase in college labor increases wages for high school graduates and therefore their contribution to public funds. In the earlier version of this paper (Findeisen and Sachs 2015a), we quantitatively elaborated the effect with a standard production function with high skill labor supply (college) and low skill labor supply (non college) in accordance with Goldin and Katz (2009). We found that the second effect even dominates the first one: general equilibrium effects seem to strengthen the argument for increasing subsidies rather than weakening it. The reason is that are approximately twice as many high-school than college graduates. Whereas these arguments where about the general desirability of increasing college financial aid, there is no reason to assume that the progressivity result is altered by general equilibrium effects on wages.

Another issue with general equilibrium is of course that individual college decisions might change in the presence of general-equilibrium effects. This is an important long-run question that is discussed in Abbott, Gallipoli, Meghir, and Violante (2016). Whereas these effects can alter the desirability of increasing financial aid in general, they should have no strong effects on the progressive nature of optimal financial aid.

# 8 Conclusion

This paper has analyzed the normative question of how to design financial aid policies for students optimally. We find the very robust and maybe ex-ante surprising result that optimal financial aid policies are strongly progressive. This result holds for different social welfare functions, assumptions on credit markets for students, and how income taxes are designed. Moreover, we find that a progressive expansion in financial aid policies could be self-financing through higher tax revenue, thus, benefitting all taxpayers as well as low income students directly. Financial aid policies seem to be a rare case with no classical equity-efficiency trade-off because a cost-effective targeting of financial goes hand in hand with goals of social mobility and redistribution. 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.

Finally, we think that our insights about tagging along the extensive margin can be applied to other important policy questions. Childcare subsidies would be an obvious application. Childcare subsidies also lead higher tax revenue through the implied increase in labor force participation. Increasing subsidies for a group of parents, where labor market participation is particularly low should therefore be more cost efficient.

# References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2016): "Education Policy and Intergenerational Transfers in Equilibrium," *Cowles Foundation Discussion Paper No. 1887R.*

Abramitzky, R., and V. Lavy (2014): "How responsive is investment in schooling to changes in redistributive policies and in returns?," *Econometrica*, 82(4), 1241–1272.

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.

Altonji, J., and T. Dunn (1996): "The Effects of Family Characteristics on the Return to Education," *The Review of Economics and Statistics*, 78(4), 692–704.

Angrist, J., D. Autor, S. Hudson, and A. Pallais (2015): "Leveling Up: Early Results from a Randomized Evaluation of Post-Secondary Aid," *Working Paper*.

Bettinger, E. (2004): "How financial aid affects persistence," in *College choices: The economics of where to go, when to go, and how to pay for it*, pp. 207–238. University of Chicago Press.

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 (2003): " 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 (2015): "Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation," *Journal of Labor Economics (forthcoming)*.

Chetty, R. (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annu. Rev. Econ.*, 1(1), 451–488.

Chetty, R., and A. Finkelstein (2013): "Social Insurance: Connecting Theory to Data," *Handbook of Public Economics*, 5, 111.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014): "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, 104(9), 2633–79.

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.

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014): "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility," *American Economic Review*, 104(5), 141–47.

Choné, P., and G. Laroque (2011): "Optimal Taxation in the Extensive Model," *Journal of Economic Theory*, 146(2), 425–453.

College Board (2015): "Trends in Student Aid 2015," *College Board*.

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. (1980): "Income Taxation with Fixed Hours of Work," *Journal of Public Economics*, 13(1), 101–110.

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 (2015a): "Designing Efficient College and Tax Policies," *Working Paper*.

Findeisen, S., and D. Sachs (2015b): "Redistribution and Insurance with Simple Tax Instruments," *Working Paper*.

——— (2016): "Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans," *Journal of Public Economics*, pp. 1–21.

Gelber, A., and M. Weinzierl (2016): "Equalizing Outcomes vs. Equalizing Opportunities: Optimal Taxation when Children's Abilities Depend on Parents' Resources," *National Tax Journal*, 69, 11–40.

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*.

Guner, N., R. Kaygusuz, and G. Ventura (2014): "Income taxation of US households: Facts and parametric estimates," *Review of Economic Dynamics*, 17(4), 559–581.

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.

Hendren, N. (2014): "The inequality deflator: Interpersonal comparisons without a social welfare function," *NBER Working Paper 20351*.

——— (2015): "The Policy Elasticity," in *Tax Policy and the Economy, Volume 30*. University of Chicago Press.

Hilger, N. G. (2015): "Parental credit constraints and college attendance: Evidence from timing of parental layoffs," *American Economic Journal: Applied Economics (forthcoming)*.

Jacobs, B. (2007): "Optimal Redistributive Tax and Education Policies in General Equilibrium," *Discussion Paper, CESifo Area Conference Public Sector Economics*.

Jacquet, L., and E. Lehmann (2016): "Optimal Taxation with Heterogeneous Skills and Elasticities: Structural and Sufficient Statistics Approaches," *THEMA Working Paper 2016-04*.

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.

Kleven, H., C. Kreiner, and E. Saez (2009): "The Optimal Income Taxation of Couples," *Econometrica*, 77(2), 537–560.

Kline, P., and C. Walters (2016): "Evaluating public programs with close substitutes: The case of Head Start," *Quarterly Journal of Economics (forthcoming)*.

Krueger, D., and A. Ludwig (2013): "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.

——— (2016): "On the Optimal Provision of Social Insurance. Progressive Taxation versus Education Subsidies in General Equilibrium," *Journal of Monetary Economics (forthcomig)*.

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

Lawson, N. (2016): "Liquidity Constraints, Fiscal Externalities and Optimal 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)*.

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.

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*.

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.*

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.

Saez, E., and S. Stantcheva (2016): "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, 106(1), 24–45.

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

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.*

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. (2015): "Optimal Income, Education, and Bequest Taxes in an Intergenerational Model," *NBER Working Paper No. 21177.*

——— (2016): "Optimal Taxation and Human Capital Policies Over the Lifecycle," *Journal of Political Economy (forthcoming).*

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

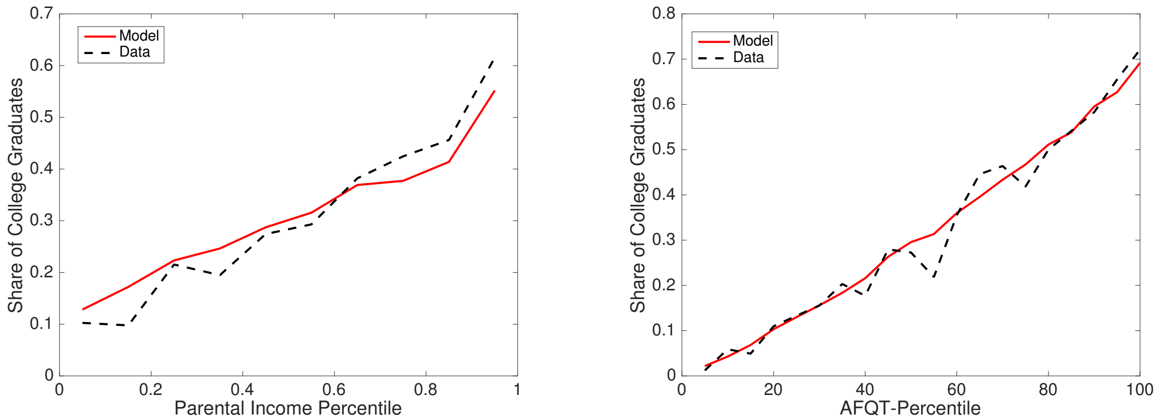
Weinzierl, M. (2014): "The Promise of Positive Optimal Taxation," *Journal of Public Economics (forthcoming).*

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.

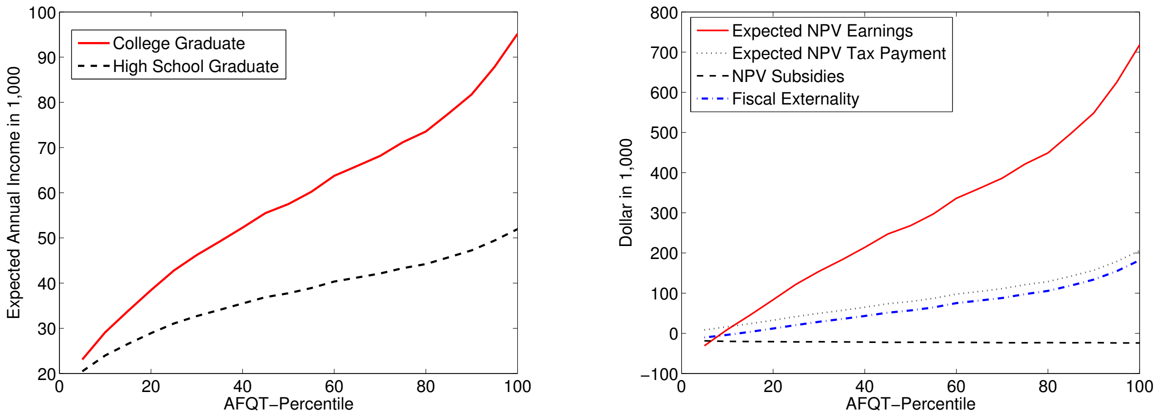
# A Online Appendix

## A.1 Additional Graphs




(a) Graduation Rates and Parental Income (b) Graduation Rates and AFQT

Figure 10: Graduation Rates

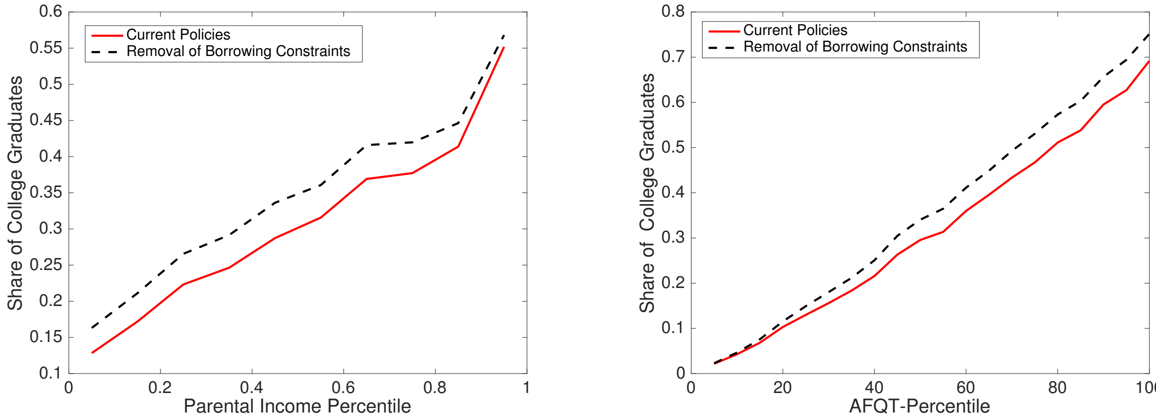



(a) Expected Annual Income (b) NPV Income and Fiscal Externality

Figure 11: Returns to College

## A.2 Derivation of Optimal Tax Formula

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 13 illustrates such a tax reform, where the marginal tax is increased by an infinitesimal amount $dT'$ in an income interval of infinitesimal length $[y(w^*), y(w^*) + dy]$.




(a) Borrowing Constraints and Parental Income (b) Borrowing Constraints and AFQT

Figure 12: Removing Borrowing Constraints

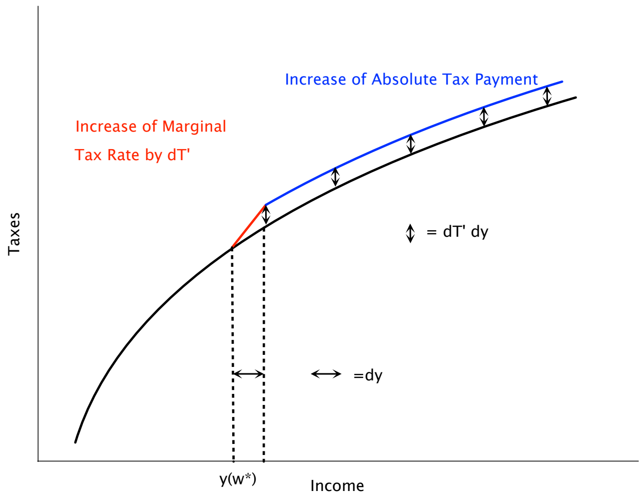



Figure 13: Tax Reform

As a consequence of this reform, all individuals with $y > y(w^*)$ (and therefore $w > w^*$) face an increase of the absolute tax level of $dT'dy$. The tax reform therefore induces a mechanical increase in tax revenue of

$$M(y(w^*)) = dT'dy \int_{y(w^*)}^{\infty} .h(y)dy$$

where

$$\begin{aligned} h(y) = & \beta^{C2} \int_{\mathbb{R}_+} \int_{\chi} \mathbb{1}_{V^C \geq V^H} k(I,X) g^C(w(y)|I,X) dIdX \\ & + \beta^H \int_{\mathbb{R}_+} \int_{\chi} \mathbb{1}_{V^C \leq V^H} k(I,X) g^C(w(y)|I,X) dIdX. \end{aligned}$$

The increase in taxes for individuals with $w > w^*$ also changes incentives for enrollment. In fact, graduation will increase by:

$$CG(y(w^*)) = dT'dy \int_{y(w^*)}^{\infty} \int_{\mathbb{R}_+} \xi(I,y)dF(I)dy.$$

where

$$\xi(I,y) = \frac{1}{f(I)} \frac{\partial F_I^C}{\partial T(y)},$$

is the semi-elasticity of college graduation with respect to an increase in $T(y)$. 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:

$$CG(y(w^*)) = dT'dy \int_{y(w^*)}^{\infty} \int_{\mathbb{R}_+} \xi(I,y)\Delta T(I,y)dF(I)dy.$$

$\Delta T(I,y)$ is the average fiscal externality of those students with parental income $y^*$ that are marginal w.r.t. a small increase in $T(y)$. It is different to (3), where the average was taken overall that are marginal wr.t. a small increase in financial aid.

In addition, an increase in the marginal tax rate also affects labor supply behavior for individuals within the interval $[y(w^*), y(w^*) + dy]$. Individuals within this infinitesimal interval change their labor supply by

$$\frac{\partial y(w^*)}{\partial \mathcal{T}'} d\mathcal{T}' = -\varepsilon_{y,1-\mathcal{T}'} \frac{y}{1-\mathcal{T}'} d\mathcal{T}'. \tag{9}$$

Whereas this change in labor supply has no first-order effect on welfare via individual utilities by the envelope theorem, it has an effect on tax revenue. The mass of these individuals is then given by

$$h(y(w^*))dy$$

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

$$LS(y(w^*)) = -\varepsilon_{y(w^*),1-\mathcal{T}'} \frac{y(w^*)}{1-\mathcal{T}'} d\mathcal{T}' h(y(w^*))dy.$$

The overall impact on welfare of the considered tax reform is thus given by

$$\Gamma(y(w^*)) = M(y(w^*)) + CG(y(w^*)) + LS(y(w^*)). \tag{10}$$

For an optimal tax system these effects have to add up to zero. $\Gamma(y(w^*)) = 0$ yields the optimal tax formula (8).

## A.3 Appendix for Section 4

### A.3.1 Current Tax Policies

We take effective marginal tax rates in the year 2000.[19] We use the year 2000 because individuals in the NLSY97 are 18 in the year 2000 on average. We set the lump sum element of the tax code $T(0)$ to minus $1,800 a year. For average incomes this fits the deduction in the US-tax code quite well.[20] For low incomes this reflects that individuals might receive transfers such as food stamps.[21] 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 and the population age structure.

### A.3.2 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 enrollment shares) tuition fee for a four-year college. We then use these numbers to

[19] 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 (2014).

[20] Guner, Kaygusuz, and Ventura (2014) 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.

[21] 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

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 A.3.5). 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.3.3 Details on Income Regressions

We first quickly explain the construction of the annuitized income variable. 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:

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

Everyone with less than 16 years of schooling is defined as a high school graduate.[22] Everyone with 16 or more years of schooling is defined as a college graduate.

We run separate regressions, one for high school graduates and one for college graduates, of the form:

$$\ln \tilde{y}_i = \alpha_{ce} + \beta_e^{IN} ln(AFQT_i) + \varepsilon_{ei}^{inc}, \tag{11}$$

[22] 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.

for $e = hs, co$. $\alpha_{ce}$ 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 (11) for each education level.

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 for all other monetary variables, incomes are measured in 2000 dollars.

Our estimates for the slopes are. $\hat{\beta}_C^{IN} = 1.88$ (0.186) and $\hat{\beta}_H^{IN} = 1.28$ (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 $\sigma^e$. We directly take the estimates for $\sigma^e$ from the distribution of residuals from (11). The values are 0.6548 for college and 0.6631 for high-school.

#### A.3.4 Parental Transfers

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,

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 $\leq$ 0.10, ** p $\leq$ 0.05, *** p $\leq$ 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:[23]

$$log(tr_i) = \alpha^{tr} + \beta_1^{tr} log(I_i) + \beta_2^{tr} co_i + \beta_3^{tr} depkids_i + \varepsilon_i^{tr}, \tag{12}$$

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. We nevertheless take this noisy measure into account because it gives our quantitative model a better fit concerning the importance of parental income. To measure how net wealth varies with parental income, we estimate the following regression:

$$w_i = \alpha^w + \beta^w I_i + \varepsilon_i^w. \tag{13}$$

We find a gradient for parental income of .127 (0.02) and an intercept of $7,950 (1164). To obtain the parental transfer for the model, we take the implied parental transfer from equation (12) and adjust it by the implied level of wealth from equation (13) and thereby recalculate the wealth into an annual transfer.

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

### A.3.5 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 grant receipt directly from the data.

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

$$gr_i = \alpha^{gr} + \beta_1^{gr} I_i + \beta_2^{gr} I_i^2 + \beta_3^{gr} black_i + \beta_4^{gr} AFQT_i + \beta_5^{gr} depkids_i + \varepsilon_i^{gr}. \tag{14}$$

Besides grant generosity being need-based (convexly decreasing) and in favor of blacks, generosity is also merit-based as $\hat{\beta}_4^{gr} > 0$ and increases with the number of other dependent children (besides the considered student) in the family.

Table 3: OLS for Grants

| | Parental Income | Parental Income$^2$ | Black | AFQT | Dependent Children |
|---|---|---|---|---|---|
| Coefficient | -.0915*** | 6.00e-07 *** | 649.06** | 23.90*** | 224.69** |
| Standard Error | (0.0192) | (1.83e-07) | (296.03 ) | ( 4.57) | (99.11) |

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

### A.3.6 Preference Estimation

Our assumptions give us a binary choice model:

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

where

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

and where $\varepsilon_i^{pc} \sim N(0, \sigma)$ as in a Probit model. 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 we have one degree of freedom in parameter choice. As expected, all the variables have a positive and significant impact on the college choice, see Table 4 for the coefficients.

Based on these estimations, we calculate the estimated psychic cost 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 $\varepsilon_i$ and then fit a normal distribution of $\kappa$ conditional on innate ability and parental income. Finally, we are then equipped with the joint distribution of parental income, innate ability and psychic costs.

Table 4: Estimation of College Graduation

| | AFQT | Father's Education | Mother's Education |
|---|---|---|---|
| Coefficient | .328*** | 2.275*** | 1.397*** |
| Standard Error | (0.034) | (0.25) | (0.28) |

N=3,897. * p $\leq$ 0.10, ** p $\leq$ 0.05, *** p $\leq$ 0.01. All coefficients multiplied by 10 000.