ONLINE APPENDIX
Optimal Need-Based Financial Aid

Mark Colas
University of Oregon

Sebastian Findeisen
University of Konstanz

Dominik Sachs
LMU Munich

May 29, 2020

In this online appendix, we present a variety of additional results, proofs, and an additional model, which is a special case of the general model and can be solved in closed-form.

Contents

1 Proofs and Background for Section 2 of Main Text

1.1 Derivation of Equation 3

. The Lagrangian for the government's problem reads as:

$$
\begin{aligned}
\mathcal{L} = & \int_{\mathbb{R}_+} \int_{\mathcal{X}} \max\{V^E(X,I), V^H(X,I)\} \tilde{k}(X,I) dX dI \\
& + \rho \Bigg\{ \int_{\mathbb{R}_+} \int_{\mathcal{X}} \mathcal{NT}^H_{NPV}(X,I) \mathbb{1}_{V_j^E < V_j^H} k(X,I) dX dI \\
& + \int_{\mathbb{R}_+} \int_{\mathcal{X}} \mathcal{NT}^G_{NPV}(X,I) \mathbb{1}_{V_j^E \geq V_j^H} P^C(X,I,\mathcal{G}(I)) k(X,I) dX dI \\
& + \int_{\mathbb{R}_+} \int_{\mathcal{X}} \mathcal{NT}^D_{NPV}(X,I) \mathbb{1}_{V_j^E \geq V_j^H} \left(1 - P^C(X,I,\mathcal{G}(I))\right) k(X,I) dX dI - \bar{F} \Bigg\}.
\end{aligned}
$$

The derivative w.r.t. $\mathcal{G}(I)$ is given by:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \mathcal{G}(I)} = & \int_{\mathcal{X}} \mathbb{1}_{V_j^E \geq V_j^H} \frac{\partial V^E(X,I)}{\partial \mathcal{G}(I)} \tilde{h}(X|I) dX \qquad (1) \\
& + \rho \int_{\mathcal{X}} \left\{ P^C(X,I,\mathcal{G}(I)) \frac{\partial \mathcal{NT}^G_{NPV}(X,I)}{\partial \mathcal{G}(I)} + \left(1 - P^C(X,I,\mathcal{G}(I))\right) \frac{\partial \mathcal{NT}^D_{NPV}(X,I)}{\partial \mathcal{G}(I)} \right\} h(X|I) dX \\
& + \rho \int_{\mathcal{X}} \mathbb{1}_{H_j \to E_j} \Big\{ P^C(X,I,\mathcal{G}(I)) \mathcal{NT}^G_{NPV}(X,I) + \left(1 - P^C(X,I,\mathcal{G}(I))\right) \mathcal{NT}^D_{NPV}(X,I) \\
& \qquad - \mathcal{NT}^H_{NPV}(X,I) \Big\} h(X|I) dX \\
& + \rho \int_{\mathcal{X}} \frac{\partial P^C(X,I,\mathcal{G}(I))}{\partial \mathcal{G}(I)} \left(\mathcal{NT}^G_{NPV}(X,I) - \mathcal{NT}^D_{NPV}(X,I) \right) h(X|I) dX
\end{aligned}
$$

Recall that $\mathbb{1}_{H_j \to E_j}$ takes the value one if an individual of type j is pushed over the college enrollment margin due to a small increase in financial aid.

The first term captures the direct utility increase of inframarginal enrollees due to receiving more financial aid. The second term captures the direct fiscal effect of paying more financial aid to inframarginal students. The third term captures the fiscal effect of additional enrollees. The fourth effect captures the fiscal effect due to the increase in the completion rate of inframarginal students. The implied change in the enrollment and dropout rate has no direct first-order effect on welfare: individuals that are marginal in their decision to enroll or not and to continue studying or drop out, were just indifferent between the two respective options, hence this change in behavior has no effect on their utility.

The definitions of $E(I)$ and $\Delta T^E(I)$ directly imply that the third term equals the enrollment effect multiplied by ρ. The definitions of $\Delta \mathcal{T}^C(I)$, $E(I)$ and $C(I)$ directly imply that the fourth term equals the completion effect multiplied by ρ.

Now it remains to be shown that the first and second term are equal to the mechanical effect. The application of the envelope theorem implies that the first term reads as

$$\int_{\chi} \mathbb{1}_{V_j^E \geq V_j^H} \mathbb{E}\left[\sum_{t=1}^{t_g^{max}} \beta^{t-1} U_c^E(\cdot)\left(1+\frac{\partial tr_t^E(\cdot)}{\partial \mathcal{G}(I)}\right) \prod_{s=1}^{t} (\mathbb{1}_{V_s^{ND} \geq V_s^D}) \prod_{s=1}^{t-1} \left(1 - Pr_s^{Grad}(X)\right)\right] \tilde{h}(X|I) dX \tilde{E}(I). \tag{2}$$

The second term, using the definitions of $\mathcal{NT}_{NPV}^{G}(X, I)$ and $\mathcal{NT}_{NPV}^{D}(X, I)$, can be written as

$$-\rho \int_{\chi} \sum_{t=1}^{t_g^{max}} \frac{1}{1+r} \prod_{s=1}^{t} P_t^E(X, I, \mathcal{G}(I)) h(X|I) dX. \tag{3}$$

Adding (2) and (3), using the definition of the social marginal welfare weight yields equation 3 form the main text.

1.2 More General Version of Equation 3 with Annual Dropout Decisions

We now show the generalization in which individuals can drop out each period. For this case, we have to distinguish between individuals that drop out in different periods. Hence, for the education decision we have: $e \in \{H, G, D_1, D_2, ..., D_{t_g^{max}}\}$, where D_t implies that individuals drop out at the beginning of year t. Accordingly we can define the net fiscal contribution of an individual of type (X, I) that drops out in period t by $\mathcal{NT}_{NPV}^{Dt}(X, I)$:

$$\mathcal{NT}_{NPV}^{Dt}(X, I) = \sum_{s=t}^{T} \left(\frac{1}{1+r}\right)^{s-1} \mathbb{E}\left(\mathcal{T}(y_s)|X, I, D_t\right) - \mathcal{G}(I) \sum_{s=1}^{t-1} \left(\frac{1}{1+r}\right)^{s-1}.$$

We also have to define the net fiscal contribution of an individual that is enrolled in year t_g^{max}

$$\mathcal{NT}_{NPV}^{Et_g^{max}}(X, I) = P_{t_g^{max}}(X, I, \mathcal{G}(I)) \mathcal{NT}_{NPV}^{G}(X, I) + \left(1 - P_{t_g^{max}}(X, I, \mathcal{G}(I))\right) \mathcal{NT}_{NPV}^{Dt_g^{max}}(X, I)$$

and for $t = 2, 3, ..., t_g^{max} - 1$:

$$\mathcal{NT}_{NPV}^{Et}(X, I) = P_t(X, I, \mathcal{G}(I)) \mathcal{NT}_{NPV}^{Et+1}(X, I) + (1 - P_t(X, I, \mathcal{G}(I))) \mathcal{NT}_{NPV}^{Dt}(X, I).$$

The Lagrangian for the government's problem reads as:

$$
\begin{aligned}
\mathcal{L} = &\int_{\mathbb{R}_+}\int_{\chi} \max\{V^E(X,I), V^H(X,I)\}\tilde{k}(X,I)dXdI \\
&+ \rho\Bigg\{\int_{\mathbb{R}_+}\int_{\chi} \mathcal{NT}^H_{NPV}(X,I)\mathbb{1}_{V_j^E<V_j^H}k(X,I)dXdI \\
&+ \int_{\mathbb{R}_+}\int_{\chi} \mathcal{NT}^G_{NPV}(X,I)\mathbb{1}_{V_j^E\geq V_j^H}P^C(X,I,\mathcal{G}(I))k(X,I)dXdI \\
&+ \sum_{\tau=1}^{t_g^{max}}\left[\int_{\mathbb{R}_+}\int_{\chi} \mathcal{NT}^{D\tau}_{NPV}(X,I)\mathbb{1}_{V_j^E\geq V_j^H}\prod_{t=1}^{\tau-1}P_t(X,I,\mathcal{G}(I))\left(1-P_\tau(X,I,\mathcal{G}(I))\right)k(X,I)dXdI\right] - \bar{F}\Bigg\}.
\end{aligned}
$$

The FOC for $\mathcal{G}(I)$ shares the same basic structure as (1). However, here the fiscal effects due to change in dropout behavior are more involved:[1]

$$
\begin{aligned}
&\rho\int_{\chi}\prod_{t=1}^{t_g^{max}-1}P_t(X,I,\mathcal{G}(I))\frac{\partial P_{t_g^{max}}(X,I,\mathcal{G}(I))}{\partial\mathcal{G}(I)}\left(\mathcal{NT}^G_{NPV}(X,I)-\mathcal{NT}^{Dt_g^{max}}_{NPV}(X,I)\right)h(X|I)dX \\
+&\rho\sum_{\tau=1}^{t_g^{max}-1}\left[\int_{\chi}\prod_{t=1}^{\tau-1}P_t(X,I,\mathcal{G}(I))\frac{\partial P_\tau(X,I,\mathcal{G}(I))}{\partial\mathcal{G}(I)}\left(\mathcal{NT}^{E\tau+1}_{NPV}(X,I)-\mathcal{NT}^{D\tau}_{NPV}(X,I)\right)h(X|I)dX\right],
\end{aligned}
$$

where we let $P_0(X,I,\mathcal{G}(I))=1$.

In shorter notation, similar to that from before, we can write

$$
\sum_{t=1}^{t_g^{max}}\left.\frac{\partial Con_t(I)}{\partial\mathcal{G}(I)}\right|_{E_t(I)}\Delta\mathcal{T}^{Con,t}(I)E_t(I)
$$

where $Con_t(I)$ is the share of those enrollees with parental income I in period t, that continue studying to year $t+1$. It is defined by

$$
Con_t(I) = \frac{E_{t+1}(I)}{E_t(I)}
$$

for $t=1,2,..,t_g^{max}-1$ and

$$
Con_{t_g^{max}}(I) = \frac{\int_\chi \mathbb{1}_{V_j^E\geq V_j^H}\prod_{s=1}^{t_g^{max}}P_s(X,I,\mathcal{G}(I))h(X|I)dX}{E_{t_g^{max}}(I)}.
$$

where

$$
E_1(I) = E(I) = \int_\chi \mathbb{1}_{V_j^E\geq V_j^H}h(X|I)dX.
$$

[1] Again, changes in dropout behaviour have no direct welfare effect due to the envelope theorem.

and

$$E_t(I) = \int_{\mathcal{X}} \mathbb{1}_{V_j^E \geq V_j^H} \prod_{s=1}^{t-1} P_s(X, I, \mathcal{G}(I)) h(X|I) dX.$$

Finally, the changes in tax revenue are defined by:

$$\Delta\mathcal{T}^{Con,t}(I) = \frac{\int_{\mathcal{X}} \Delta\mathcal{T}^{Con,t}(X, I) \frac{\partial P(X,I,\mathcal{G}(I))}{\partial \mathcal{G}(I)} h(X|I) dX}{\int_{\mathcal{X}} \frac{\partial P(X,I,\mathcal{G}(I))}{\partial \mathcal{G}(I)} h(X|I) dX}$$

where

$$\Delta\mathcal{T}^{Con,t}(X, I) = \mathcal{NT}_{NPV}^{Et}(X, I) - \mathcal{NT}_{NPV}^{Dt}(X, I).$$

Hence, the equivalent to equation 3 is given by:

$$\frac{\partial E(I)}{\partial \mathcal{G}(I)} \times \Delta\mathcal{T}^E(I) + \sum_{t=1}^{t_g^{max}} E_t(I) \left.\frac{\partial Con_t(I)}{\partial \mathcal{G}(I)}\right|_{E_t(I)} \Delta\mathcal{T}^{Con,t}(I) - \tilde{E}(I)\left(1 - W^E(I)\right).$$

1.3 Optimal Financial Aid in the Structural Model

In this section we return to the the optimality condition for financial aid, and highlight which structural parameters are key for the relationship between optimal financial aid and parental income in our quantitative model. For brevity and clarity, we focus on the share of marginal and inframarginal enrollees because our numerical analysis below shows that these are the most important forces for our progressivity result.

Inframarginal Enrollees: For brevity we focus on the share of inframarginal enrollees $E(I)$ instead of $\tilde{E}(I)$.[2] It is given by:

$$E(I) = \int_{\tilde{X}} \frac{\exp\left(\tilde{V}^E(\tilde{X}, I)/\sigma^E\right)}{\exp\left(\tilde{V}^E(\tilde{X}, I)/\sigma^E\right) + \exp\left(V^H(\tilde{X}, I)/\sigma^E\right)} dH^*(\tilde{X}|I).$$

where $H^*(\tilde{X}|I)$ is the CDF for $\tilde{X}$ conditional on I and where $\tilde{V}^E\left(\tilde{X}, I\right) = V^E(X, I) - \varepsilon^E$ is the value of enrolling in college minus the idiosyncratic taste for college ε^E[3]. This expression immediately follows from the fact that the idiosyncratic enrollment benefit ε^E is distributed according to a type I extreme value distribution with scale parameter σ^E. The number of enrollees

[2] The insights would be identical if we were looking at $\tilde{E}(I)$ here but notation would be unnecessarily cumbersome.

[3] Note that $V^H(\tilde{X}, I) = V^H(X, I)$, with some abuse of notation, because the idiosyncratic preference term ε^E does not affect $V^H(X, I)$

conditional on $\left(\tilde{X}, I\right)$ increases in the difference in the value functions of attending college or not. How $E\left(I\right)$ varies with parental income is largely determined by the relation of parental income with (i) psychic costs κ, (ii) parental transfers, (iii) ability.

Marginal Enrollees: For a given $(\tilde{X}, I)$, the share of marginal enrollees is given by

$$\frac{\partial E(\tilde{X}, I)}{\partial \mathcal{G}(I)} = \frac{E(\tilde{X}, I)\left(1 - E(\tilde{X}, I)\right)}{\sigma^E} \frac{\partial V^E(\tilde{X}, I)}{\partial \mathcal{G}(I)},$$

where $E(\tilde{X}, I)$ is the enrollment share of individuals with observable characteristics $\tilde{X}$ and income I, $\frac{E(\tilde{X},I)\left(1-E(\tilde{X},I)\right)}{\sigma^E}$ is the density of the enrollment benefit parameter ε^E at the value where an $(\tilde{X}, I)$ individual is indifferent between enrolling in college or not. Formally, this threshold is given by $\tilde{\varepsilon}^E(\tilde{X}, I) = \tilde{V}^E(\tilde{X}, I) - V^H(\tilde{X}, I)$. Intuitively, the higher this density, the more individuals are marginal in their decision and the stronger is the increase in enrollment due to higher financial aid. A property of the logit distribution is that the density is maximized if enrollment is at 50%, as is the case also for a normal distribution. Further, the lower the scale parameter σ^E, the higher the share of marginal students ceteris paribus.

The share of marginal enrollees also depends on how much this threshold $\tilde{\varepsilon}^E(\tilde{X}, I)$ changes due to an increase in financial aid, which is captured by:

$$\frac{\partial V^E(\tilde{X}, I)}{\partial \mathcal{G}(I)} = \mathbb{E}\left[\sum_{t=1}^{t_g^{max}} \beta^{t-1} c_t(\cdot)^{-\gamma} \left(1 + \frac{\partial tr_t^E(\tilde{X}, I, \mathcal{G}(I))}{\partial \mathcal{G}(I)}\right) \prod_{s=1}^{t} (\mathbb{1}_{V_s^{ND} \geq V_s^D}) \prod_{s=1}^{t-1} \left(1 - Pr_s^{Grad}(X)\right)\right].$$

Intuitively, agents with low marginal utility $c_t(\cdot)^{-\gamma}$ during college react more strongly financial aid changes. According to this logic, children with low parental income should be more responsive to increases in financial aid. How much this effect varies with parental income is governed by γ, which we estimate with maximum likelihood. In addition, the stronger the crowding out of the parental transfer ($-\frac{\partial tr_t^E(\tilde{X},I,\mathcal{G}(I))}{\partial \mathcal{G}(I)}$), the less responsive are individuals ceteris paribus since less of the financial aid increase reaches them.

2 Is Optimal Financial Aid Progressive? A Simple Model

2.1 Model

Simplified Environment. We assume that preferences are linear in consumption and that labor incomes are taxed linearly at rate τ, which is larger than 0 and smaller than one. We consider a static problem. If individuals do not go to college, they earn income y_H. If they go to college, they pay tuition $\mathcal{F}$ and earn $y_H(1 + \theta)$. Individuals are heterogeneous in ability/returns to college, θ, and each $\theta > 0$. There is no uncertainty. Further, individuals are heterogeneous in parental income I. If individuals go to college, they receive a parental transfer $tr(I)$ with $tr'(I) > 0$ and financial aid $\mathcal{G}(I)$.

Individual Problem. If an individual decides against college, utility is given by $U^H = (1-\tau)y_H$. If an individual goes to college, utility is given by $U^C(\theta, I) = (1-\tau)\, y_H\, (1+\theta) - (\mathcal{F} - \mathcal{G}(I) - tr(I))$. For each income level I, we can define the ability of the marginal college graduate $\tilde{\theta}(I)$, implicitly given by $U^H = U^C(\tilde{\theta}(I), I)$. All types (θ, I) with $\theta \geq (<)\tilde{\theta}(I)$ (do not) attend college. Note that higher parental income here simply has the role of lowering the costs of college. This implies that high-parental-income children are more likely to select into college. This channel is reinforced if there is a positive association between I and θ.

Government Problem and Optimal Financial Aid for a Given I. The government uses non-negative Pareto weights over the types as in the general model from the last section. Consistent with the notation from last section, $F(I)$ is the parental income distribution and $H(\theta|I)$ the conditional distribution of ability. Appendix 2.2 shows that the following equation holds:

$$\underbrace{\frac{h(\tilde{\theta}(I)|I)}{y_H(1-\tau)}}_{\frac{\partial E(I)}{\partial \mathcal{G}(I)}} \times \underbrace{\Big(\tau y_h \tilde{\theta}(I) - \mathcal{G}(I)\Big)}_{\Delta \mathcal{T}^E(I)} - \underbrace{\Big(1 - H(\tilde{\theta}(I)|I)\Big)}_{\tilde{E}(I)}(1 - W^E(I)) = 0.$$

First note that there is no completion effect since we abstract from dropout. Second, the fiscal externality takes a simple form. Third the ratio of marginal over inframarginal students is determined by the hazard rate of the conditional skill distribution. Rewriting leads to a rather tractable expression for optimal financial aid $\mathcal{G}(I)$.

Proposition 1. *The optimal financial aid schedule in the simplified enviornment is given by*

$$\mathcal{G}(I) = \tau\,(\mathcal{F} - tr(I)) - y_H(1-\tau)^2 \frac{\Big(1 - H(\tilde{\theta}(I)|I)\Big)}{h(\tilde{\theta}(I)|I)} \times (1 - W^E(I)), \tag{4}$$

where $\tilde{\theta}(I) = \frac{\mathcal{F} - tr(I) - \mathcal{G}(I)}{(1-\tau) y_H}$. *and* $\tau\,(\mathcal{F} - tr(I)) = \tau y_H \tilde{\theta}(I)$.

Proof. See Section 2.2 of this online appendix. □

The first term in (4), $\tau\,(\mathcal{F} - tr(I))$, can be interpreted as a Pigouvian correction. Without any distortions, i.e. $\mathcal{G}(I) = \tau = 0$, the marginal college enrollee would be characterized by

$$\theta^*(I) y_H = \mathcal{F} - tr(I). \tag{5}$$

Here the private returns and costs are equalized to the social ones. Such a condition is typically called "first best". When τ or $\mathcal{G}(I) \neq 0$, the marginal enrollee still equates private returns to private costs, but there is a wedge between the *social* returns and costs now. Equating private returns and costs yields:

$$\tilde{\theta}(I)(1-\tau)y_H = \mathcal{F} - tr(I) - \mathcal{G}(I). \tag{6}$$

Comparing (6) with (5) shows that the fiscal externality $\Delta\mathcal{T}^E(I) = \tau y_h \tilde{\theta}(I) - \mathcal{G}(I)$ can be seen as a wedge. This is the classical "siamese twins" result of Bovenberg and Jacobs (2005): the sole presence of taxes gives a rationale for subsidizing education and the size of the subsidy is increasing in the size of the tax. Setting $\mathcal{G}(I) = \tau(\mathcal{F} - tr(I)) = \tau y_H \theta^*(I)$ would imply $\tilde{\theta}(I) = \theta^*(I)$ and hence yield the first-best education level. Such a schedule would be optimal if there would be no desire for redistribution ($\frac{\partial W^E(I)}{\partial I} = 0$) and if the government budget constraint would be exactly satisfied for such a subsidy schedule (which would then imply $W^E(I) = 1 \; \forall \; I$). In the more likely case where the government budget constraint is not fulfilled, then also the second terms shows up that captures the transfers that are made to the inframarginal students.

The reason is that when choosing the optimal education subsidy $\mathcal{G}(I)$, the social planner has to account for the fact that an increase in $\mathcal{G}(I)$ also has to be paid to those students that are inframarginal in their decision.[4] This is accounted for in the second part of (4). Since the decision of inframarginal students is not altered, this is a pure transfer which is valued by $W^E(I) - 1$ multiplied with the share of inframarginal students. This implies that if $W^E(I) > (<)1$, the planner would subsidize students of parental income up to a point where education is above (below) the first best level as defined above.[5] Further, this second term inversely proportional to share of marginal students. Intuitively, the more marginal students can be incentivized, the higher is the relative weight on the first term.

In the following we want to explore whether financial aid optimally decreases with parental income. For this purpose, we shut down any redistributive case for financial aid and assume that $\frac{\partial W^E(I)}{\partial I} = 0$. Two useful benchmark cases generate this: (i) a government that solely wants to maximize tax revenue (implying $W^E(I) = 0$ for all I) and (ii) unweighted Utilitarinism (implying $W^E(I) = constant < 1$ for all I (if the first-best rule is not budget-feasible) as we elaborate in Section 2.2 of this online appendix). If redistribution within college students is desired, i.e. with declining weights $W^{E\prime}(I) < 0$, this would strengthen the case for progressivity and need-based financial aid.

Is Optimal Financial Aid Decreasing in Parental Income? We proceed in two steps and first state a result on the progressivity if parental income and child's ability are independently distributed.

[4] If the planner can choose $\mathcal{G}(I,\theta)$ in this simple model, she effectively has lump-sum taxes/transfers available (for all college students). She only needs to correct the fiscal externality in this case (the other considerations like the ratio of marginal to inframarginals and redistribution within students can be perfectly dealt with by choosing $\mathcal{G}(I,\theta)$ for each type. This is not the case in the more general model presented in Section 2. We analyze the case of jointly optimizing merit-based and need-based financial aid quantitatively in 6.7.

[5] This resembles the results of the optimal income tax literature with extensive margin labor supply responses that negative participation taxes are optimal if the social welfare weight of low income workers is above one, see e.g. Saez (2002).

Corollary 1. *Assume that ability θ and parental income I are independent, that is, $H(\theta|I) = H(\theta)\ \forall\ \theta, I$. Further assume $\frac{\partial W^E(I)}{\partial I} = 0$, i.e. there is no desire to redistribute from high to low parental income students. Then the optimal financial aid schedule is progressive (i.e., $\mathcal{G}'(I) < 0\ \forall\ I$) if the distribution $H(\theta)$ is log concave.*

Proof. See Section 2.3 of this online appendix. $\square$

The first term in (4) is decreasing in I. The higher parental income, the lower are the costs of college $\mathcal{F} - tr(I)$ and hence, for a given rate of subsidization τ, the lower is the overall level of the subsidy. Since $\tilde{\theta}'(I) < 0$,[6] the second term is decreasing in I if the inverse of the hazard rate of $H(\theta)$ is decreasing. As Bagnoli and Bergstrom (2005) point out, log-concavity of a density function is sufficient for an increasing hazard rate.[7] Hence, in the illustrative case in which parental income and child's ability are independent, we have an important benchmark, where the selection mechanism through parental income in itself calls for progressive financial aid. Next we turn to the empirically more appealing case in which parental income and ability are positively associated.[8]

Corollary 2. *Assume that ability θ and parental income I are positively associated in the sense that for $I' > I$, the distribution $H(\theta|I')$ dominates $H(\theta|I)$ in the hazard rate order, that is,*

$$\forall \theta, I, I' \text{ with } I' > I: \ \frac{h(\theta|I)}{1 - H(\theta|I)} \geq \frac{h(\theta|I')}{1 - H(\theta|I')}. \tag{7}$$

Further assume $\frac{\partial W^E(I)}{\partial I} = 0$, i.e. there is no desire to redistribute from high to low parental income students. Then the optimal financial aid schedule is progressive (i.e. $\mathcal{G}'(I) < 0\ \forall\ I$) if the conditional skill distributions $H(\theta|I)$ are log concave.

Proof. See Section 2.4 of this online appendix. $\square$

This condition (7) is stronger than first-order stochastic dominance (FOSD) but does imply that the skill distribution of higher parental income levels first-order stochastically dominates the skill distribution of lower parental income levels. FOSD of the skill distribution, however, does not automatically imply (7).[9] For the empirically plausible Pareto distribution, FOSD does imply dominance in the hazard rate order. Consider, for example, the specification $h(\theta|I) = \alpha(I)\frac{\theta^{\alpha(I)}}{\theta^{\alpha(I)+1}}$, where $\alpha(I)$ is the thickness parameter. Here we have $\frac{1-H(\theta|I)}{h(\theta|I)} = \frac{\theta}{\alpha(I)}$ and hence if $\alpha'(I) < 0$, then

[6] Note that for this we need $tr'(I) + \mathcal{G}'(I) > 0$, i.e. that financial aid is not too progressive. As our proof in Appendix 2.3 shows, this is the case.

[7] Log-concavity of a probability distribution is a frequent condition used in many mechanism design or contract theory applications, as this is "just enough special structure to yield a workable theory" (Bagnoli and Bergstrom, 2005).

[8] As Carneiro and Heckman (2003, p.27) write: "Family income and child ability are positively correlated, so one would expect higher returns to schooling for children of high income families for this reason alone." In a famous paper, Altonji and Dunn (1996) find higher returns to schooling for children with more-educated parents than for children with less-educated parents.

[9] See, e.g., Shaked and Shanthikumar (2007, p.18).

the tail of the skill distribution of high-parental-income children is thicker and the FOSD property is fulfilled. Therefore, (7) is fulfilled.

The goal of this section was to show that under some rather weak assumptions, optimal financial aid is indeed decreasing in income.

2.2 Proof of Proposition 1

The government's problem reads as

$$\begin{aligned}
\max_{\mathcal{G}(I)} & \int_{\mathbb{R}_+} \int_{\underline{\theta}}^{\tilde{\theta}(I)} U\left((1-\tau) y_H\right)\, d\tilde{H}(\theta|I) d\tilde{F}(I) \\
& + \int_{\mathbb{R}_+} \int_{\tilde{\theta}(I)}^{\overline{\theta}} U\left(\left(\left(1-\tau y_H\right)(1+\theta) - \left(\mathcal{F} - \mathcal{G}(I) - tr(I)\right)\right)\right) d\tilde{H}(\theta|I) d\tilde{F}(I) \\
& + \lambda \left[\int_{\mathbb{R}_+} \int_{\underline{\theta}}^{\tilde{\theta}(I)} \tau y_H \, dH(\theta|I) dF(I) + \int_{\mathbb{R}_+} \int_{\tilde{\theta}(I)}^{\overline{\theta}} \left(\tau y_H (1+\theta) - \mathcal{G}(I)\right)\, dH(\theta|I) dF(I) - \bar{F} \right],
\end{aligned}$$

where, as in Section 2, $\tilde{H}$ and $\tilde{F}$ denote the distributions of Pareto weights which integrate up to one and $\bar{F}$ is some exogenous revenue requirement. All Pareto weights are non-negative.

The first-order condition for $\mathcal{G}(I)$ is given by:

$$h(\tilde{\theta}(I)|I) \left| \frac{\partial \tilde{\theta}(I)}{\partial \mathcal{G}(I)} \right| \left(\tau y_h \tilde{\theta}(I) - \mathcal{G}(I) \right) - \left(1 - H(\tilde{\theta}(I)|I) \right) (1 - W^E(I)) = 0, \tag{8}$$

where $W^E(I)$ is average social marginal welfare weight for enrollees with parental income I and formally given by:

$$W^E(I) = \frac{\int_{\tilde{\theta}}^{\overline{\theta}} U'\left(\left(\left(1-\tau y_H\right)(1+\theta) - \left(\mathcal{F} - \mathcal{G}(I) - tr(I)\right)\right)\right) d\tilde{H}(\theta|I) \tilde{f}(I)}{\lambda(1 - H(\theta|I)) f(\theta)}.$$

Using

$$\frac{\partial \tilde{\theta}(I)}{\partial \mathcal{G}(I)} = -\frac{1}{y_H(1-\tau)}$$

and inserting into (8) gives the first-order condition explained in the main text just before Proposition 1. Solving for $\mathcal{G}(I)$ gives Proposition 1.

Unweighted Utilitarianism with U(x)=x. If we assume $\tilde{h}(\theta|I) = h(\theta|I)$ for all (θ, I) and further assume $U(x) = x$, then, we obtain

$$W^E(I) = \frac{1}{\lambda} \; \forall \; I,$$

hence the welfare weights are the same for all parental income groups. The value for λ is measures the marginal value of public funds and therefore depends on $\bar{F}$. To get an understanding of the optimal value for λ, consider the perturbation around the optimum, where $\mathcal{G}(I)$ is increased by $\delta \to 0$ for all I. This basically implies changing the lump sum component of $\mathcal{G}(I)$ and is equivalent to just integrating over (8). The impact on welfare is given by:

$$\int_I (1 - W^E(I)) \left(1 - H(\tilde{\theta}(I)|I)\right) dF(I) + \int_I h(\tilde{\theta}(I)|I) \left|\frac{\partial \tilde{\theta}(I)}{\partial \mathcal{G}(I)}\right| \left(\tau y_h \tilde{\theta}(I) - \mathcal{G}(I)\right) dF(I) = 0.$$

and hence for $W^E(I) = \frac{1}{\lambda} \; \forall \; I$:

$$\int_I \left(1 - \frac{1}{\lambda}\right) \left(1 - H(\tilde{\theta}(I)|I)\right) dF(I) + \int_I h(\tilde{\theta}(I)|I) \left|\frac{\partial \tilde{\theta}(I)}{\partial \mathcal{G}(I)}\right| \left(\tau y_h \tilde{\theta}(I) - \mathcal{G}(I)\right) dF(I) = 0.$$

Here we see that $\lambda = 1$ would be consistent with $\mathcal{G}(I) = \tau y_h \tilde{\theta}(I)$ for all I. Recall that the government budget constraint is given by:

$$\int_{\mathbb{R}_+} \int_{\underline{\theta}}^{\tilde{\theta}(I)} \tau y_H \; dH(\theta|I) dF(I) + \int_{\mathbb{R}_+} \int_{\tilde{\theta}(I)}^{\bar{\theta}} \left(\tau y_H \left(1 + \theta\right) - \mathcal{G}(I)\right) \; dH(\theta|I) dF(I) - \bar{F} = 0.$$

If the exogenous revenue requirement $\bar{F}$ is such that the budget constraint holds for $\mathcal{G}(I) = \tau y_h \tilde{\theta}(I)$ for all I, then we obtain $\lambda = 1$ and the formula in Proposition 1 becomes

$$\mathcal{G}(I) = \tau \left(\mathcal{F} - tr(I)\right). \tag{9}$$

Assume that instead the budget constraint would be violated and this level of financial aid can not be financed. Then we have $\lambda > 1$ and hence $W^E(I) < 1 \; \forall \; I$. Generally, we could also have the case where $\lambda < 1$. E.g. assume that $\bar{F} = -\infty$. In this case, there would be infinitely many public funds available for financial aid and therefore the marginal value of public funds would be zero. But of course this is only of theoretical interest.

The fact that the marginal value of public funds is not equal to unity even though preferences are linear may seem in contrast to the optimal income tax literature, where it is a standard result that the marginal value of public funds is equal to one for quasi-linear preferences and in other words the average welfare weights is equal to one, see e.g. Saez (2002). The reason is that the policy instruments that we consider are such that there is no lump sum element. While the financial aid schedule $\mathcal{G}(I)$ of course has an intercept $\mathcal{G}(0)$ that can optimally be chosen, this is no lump sum transfer in the classical sense because it only reaches college students and not individuals who forgo college. Therefore, varying this lump sum component also has incentive effects on the college decision and one cannot just pay out a dollar to everyone without affecting behavior.

2.3 Proof of Corollary 1

Differentiating (4) w.r.t. I yields:

$$\mathcal{G}'(I) = -\tau tr'(I) + (1-\tau)\frac{\partial\left(\frac{1-H(\tilde{\theta}(I))}{h(\tilde{\theta}(I))}\right)}{\partial\tilde{\theta}(I)}\left(tr'(I)+\mathcal{G}'(I)\right)(1-W^E)$$

where we used $\tilde{\theta}(I) = \frac{\mathcal{F}-tr(I)-\mathcal{G}(I)}{(1-\tau)y_H}$. and therefore $\tilde{\theta}'(I) = \frac{-tr'(I)-\mathcal{G}'(I)}{(1-\tau)y_H}$. Solving for $\mathcal{G}'(I)$ we get

$$\mathcal{G}'(I) = \frac{-\tau tr'(I) + (1-\tau)\frac{\partial\left(\frac{1-H(\tilde{\theta}(I))}{h(\tilde{\theta}(I))}\right)}{\partial\tilde{\theta}(I)}tr'(I)(1-W^E)}{1-(1-\tau)\frac{\partial\left(\frac{1-H(\tilde{\theta}(I))}{h(\tilde{\theta}(I))}\right)}{\partial\tilde{\theta}(I)}(1-W^E)}$$

which proves Corollary 1 since by assumption $tr'(I)>0$ and log concavity of the skill distribution implies $\frac{\partial\left(\frac{1-H(\tilde{\theta}(I))}{h(\tilde{\theta}(I))}\right)}{\partial\tilde{\theta}(I)} < 0$.

2.4 Proof of Corollary 2

Differentiating (4) w.r.t. I yields:

$$\mathcal{G}'(I) = -\tau tr'(I) + (1-W^E)(1-\tau)\left[\frac{\partial\left(\frac{1-H(\tilde{\theta}(I)|I)}{h(\tilde{\theta}(I)|I)}\right)}{\partial\tilde{\theta}(I)}\left(tr'(I)+\mathcal{G}'(I)\right) - y_H(1-\tau)\left.\frac{\partial\left(\frac{1-H(\theta|I)}{h(\theta|I)}\right)}{\partial I}\right|_{\theta=\tilde{\theta}(I)}\right]$$

Hence we obtain

$$\mathcal{G}'(I) = \frac{-\tau tr'(I) + (1-W^E)(1-\tau)\left[\frac{\partial\left(\frac{1-H(\tilde{\theta}(I))}{h(\tilde{\theta}(I))}\right)}{\partial\tilde{\theta}(I)}tr'(I) - y_H(1-\tau)\left.\frac{\partial\left(\frac{1-H(\theta|I)}{h(\theta|I)}\right)}{\partial I}\right|_{\theta=\tilde{\theta}(I)}\right]}{1-(1-\tau)\frac{\partial\left(\frac{1-H(\tilde{\theta}(I))}{h(\tilde{\theta}(I))}\right)}{\partial\tilde{\theta}(I)}(1-W^E)}$$

which proves Corollary 2 since by assumption $tr'(I)>0$, log concavity of the skill distribution implies $\frac{\partial\left(\frac{1-H(\tilde{\theta}(I))}{h(\tilde{\theta}(I))}\right)}{\partial\tilde{\theta}(I)} < 0$ and we assumed

$$\frac{\partial\left(\frac{1-H(\theta|I)}{h(\theta|I)}\right)}{\partial I} > 0 \;\forall\, \theta, I.$$

3 Estimation and Calibration

3.1 Current Tax Policies and Tuition

To capture current tax policies, we use the approximation of Heathcote et al. (2017), which has been shown to work well in replicating the US tax code. Since this specification does not contain a lump-sum element, we slightly adjust this schedule. 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.[10] For low incomes this reflects that individuals might receive transfers such as food stamps.[11]

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. We also take into account the amount of money that is spent per student by public appropriations, which has to be taken into account for the fiscal externality. The average values are \$7,434 for annual tuition and \$4,157 for annual public appropriations per student. Besides these implicit subsidies, students receive explicit subsidies in the form of grants and tuition waivers. We estimate how this grant receipt varies with parental income and ability in Appendix 3.2 using information provided in the NLSY97. We find a strong negative effect of parental income on financial aid receipt. Additionally, we can capture merit-based grants by the conditional correlation of AFQT scores with grant receipt. Finally, we calibrate the exogenous budget element $\bar{F}$ in the following way. For the current U.S. polices, we calculate the present value of financial aid spending and the present value of tax revenues collected from the cohorts that we consider (born between 1980 and 1984 from the NLSY97) and obtain $\bar{F}$ from the difference between the two.

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

[10] Guner et al. (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.

[11] 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

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 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 (North East), \$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 3.2). 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).

3.2 Estimation of Grant Receipt

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 as a Tobit model:

$$gr_i = \alpha^{gr} + f(I_i) + \beta_4^{gr} AFQT_i + \beta_5^{gr} depkids_i + \varepsilon_i^{gr}. \tag{10}$$

where $f(I_i)$ is a spline function of parental income and ε_i^{gr} represents measurement error. Besides grant generosity being need-based (convexly decreasing), 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 1: OLS for Grants

	AFQT	Dependent Children
Coefficient	39.40***	321.75**
Standard Error	(5.03)	(106.39)

N=968. * p $\leq$ 0.10, ** p $\leq$ 0.05, *** p $\leq$ 0.01.

3.3 Wage Estimation

We specify and estimate wage life-cycle paths as follows. Our procedure first estimates labor earnings life-cycle profiles and then calibrates the respective wage profiles based on those estimates in a second step. Specifically, we use the following functional form for earnings y :

$$\forall\ e = H, G:\ \log y_{it}^{e} = \beta_0^{es} + \beta_\theta^e \log \theta_i + \beta_{t1}^e t + \beta_{t2}^e t^2 + \beta_{t3}^e t^3 + v_i^{e*}. \tag{11}$$

We estimate separate parameters for high school graduates and college graduates.[12] The parameter β_θ^e captures different returns to ability for agents of a given education level. The extent to which the college wage premium is increasing in ability is determined by the ratio $\frac{\beta_\theta^G}{\beta_\theta^H}$. We find a ratio larger than 1, which implies a complementary relationship between initial ability and education. Our estimates can be found in Table 2. $v_i^{e^*}$ is a random effect that captures persistent differences in wages conditional on the agent's schooling choice. We assume that agents do not know the value of $v_i^{e^*}$ at the beginning of the model, but that its value is revealed as soon as the agents finish their education and enter the labor market. Uncertainty over $v_i^{e^*}$ creates uncertainty over an agent's returns to college. After $v_i^{e^*}$ there is no further uncertainty about an agent's wage path.

The age earnings coefficients $\beta_{t1}^e, \beta_{t2}^e$ and β_{t3}^e are education dependent but independent from gender. However, since we assume different labor supply elasticities for men and women, the implied wage life-cycle profiles will differ across gender because how a given earnings path maps into wages depends on the labor supply elasticity. The age coefficients are estimated from the NLSY79 since individuals from the NLSY97 are only observed until their mid-30s. In sum, this procedure pins down a stochastic distribution of potential life-cycle wage paths for each individual, which depend on gender, ability, and the education decisions.[13]

We estimate the age coefficients β_{t1}^e, β_{t2}^e , β_{t3}^e using panel data from the NLSY79 since individuals in the NSLY97 are too young (born between 1980 and 1984) such that we can infer how wages evolve once individuals are older than 35. In the second step, we build the transformed variable $\widetilde{\log y_{it}^e} = \log y_{it}^e - \beta_t^e t - \beta_{t2}^e t^2 - \beta_{t3}^e t^3$, which takes out age affects from yearly log incomes. Using the NLSY97, we estimate the relationship of log income with gender and log AFQT, esti-

[12] Dropouts have the same wage parameters as high school graduates except for the constant term. This gives us a very good fit for the relative earnings of dropouts, consistent with the evidence in Lee et al. (2017).

[13] We use these same parameter estimates to calculate life-cycle earnings for parents. We choose the idiosyncratic competent of earnings, v_i^{e*}, to generate earnings at age 45 equal to the parental earnings levels we observe in the data.

mating separate models and coefficients by education level. We use a random-effects estimator and assume normality, yielding education specific variances for v_i^e. The estimates are displayed in Table 2. There is a significant college premium in the model, although the high-school constant is larger, because we have used education dependent age profiles.

College Educated				
	Female	Log AFQT	Education Constant	Variance v_i
Coefficient	-0.14***	0.47***	3.06***	0.42
Standard Error	(0.02)	(0.07)	(0.35)	

High-School Educated				
	Female	Log AFQT	Education Constant	Variance v_i
Coefficient	-0.25***	0.31***	7.11***	0.36
Standard Error	(0.01)	(0.03)	(0.35)	

Table 2: Regressions: Income

Notes: Random effect models, estimated with NLSY9. Dependent variable is log yearly income, cleaned for age effects. Age effects are obtained by estimating a cubic polynomial on the NLSY79. These age coefficients are available upon request. N=10,165 (College) and N=19,955 (High-School) . * p $\leq$ 0.10, ** p $\leq$ 0.05, *** p $\leq$ 0.01.

Next, we explain how to go from the estimated income to the wage profiles. The reason why we do not estimate wage profiles directly is that we append Pareto tails to the income distribution on which more reliable information is available. 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 \$150,000. We set the shape parameter α of the Pareto distribution to 1.5 for all income distributions.

Next we describe the mapping from y to w as in Saez (2001). Given the utility function we assume with no income effects, in each year individuals solve a static labor supply problem where optimal labor supply in that year only depends on the current wage (which evolves over the life-cycle) and marginal tax distortions. It is easy to show that the first-order condition for an individual facing a marginal tax rate schedule is

$$\ln w = \frac{\epsilon + \tau}{1 + \epsilon} \ln y - \frac{1}{1 + \epsilon} \ln(\lambda\,(1 - \tau)),$$

if the tax function is of the form $T(y) = y - \rho y^{1-\tau}$. Using the estimates from the regression model, we can express the wage for a given type (age, gender, ability, education) as at age t:

$$\ln w_{it} = \frac{\epsilon + \tau}{1+\epsilon} \left(\hat{\beta}_0^{es} + \hat{\beta}_\theta^e \log \theta_i + \hat{\beta}_t^e t + \hat{\beta}_{t2}^e t^2 + \hat{\beta}_{t3}^e t^3 + v_i^{e^*} \right) - \frac{1}{1+\epsilon} \ln(\lambda (1-\tau)).$$

Parent's Earnings Profile Calibration We assume that parental earnings are determined by a similar process to the child's earnings. Specifically, parental earnings are given by

$$\forall\, e = H, G:\ \log y_t^P = \beta_{t1}^{\text{ParEdu}} \text{ParAge}_t + \beta_{t2}^{\text{ParEdu}} \text{ParAge}_t^2 + \beta_{t3}^{\text{ParEdu}} \text{ParAge}_t^3 + v^P.$$

where ParAge_t is the parent's age in period t. The age coefficients, $\beta_{t1}^{\text{ParEdu}}$, $\beta_{t2}^{\text{ParEdu}}$, and $\beta_{t3}^{\text{ParEdu}}$ are taken from the child's earnings regression. We assume that the parent's age coefficients are given by the college age coefficients if at least one parent has attended college, otherwise the parent's age coefficients are given by the age coefficients for a child that has not attended college.

The term v^P represents persistent, idiosyncratic differences in earnings across parents. We assume that we observe the parental income variable I when parents are 40 years old. Therefore, we must have $y_{40} = I$ for each parent we observe in data. We therefore choose v^P such that the predicted parental income at age 40 is equal to the observed parental income variable I. We can write this as

$$v^P = \log I - \left(\beta_{t1}^{\text{ParEdu}} \text{ParAge}_t + \beta_{t2}^{\text{ParEdu}} \text{ParAge}_t^2 + \beta_{t3}^{\text{ParEdu}} \text{ParAge}_t^3 \right).$$

3.4 Likelihood Function

Assume that the econometrician observes transfers $tr_i^{e,o}$, which differ from true transfers, $tr_i^{e\star}$, by an error term e^{tr}. Further, we assume this error term is normally distributed: $e^{tr} \sim N\left(0, \sigma^{e^{tr}}\right)$. We suppress all dependencies for notational convenience. Then, given parameters Γ, the likelihood contribution of an agent who graduates from college after T_i^E years, has a sequence of work in college decisions of $\left\{\ell_{it}^E\right\}_{t=1}^{T_i^E}$, and has observed college transfers $tr_i^{E,o}$ is[14]

$$\begin{aligned} &\mathcal{L}_i \left(e_i = G, tr_i^{E,o}, \left\{\ell_{it}^E\right\}_{t=1}^{T_i^E} | \Gamma \right) = \\ &Pr\left(E\right) f^N \left(\frac{tr_i^{E\star} - tr_i^{E,o}}{\sigma^{e^{tr}}} \right) \frac{1}{\sigma^{e^{tr}}} \left(\prod_{t=1}^{T_i^E} Pr\left(\ell_{it}^E\right) \right), \end{aligned} \tag{12}$$

where f^N is the standard normal PDF, and where the probability of initially enrolling in college, $Pr\left(E\right)$, and the choice probability of not dropping out and working ℓ_{it}^E in college, $Pr\left(\ell_{it}^E\right)$, are given by the extreme-value choice probabilities as

[14]The probability of this event in fact also depends on the graduation probabilities Pr_t^{Grad}. But these are just constant factors in the likelihood, which is why refrain from putting them here.

$$Pr\left(E\right) = \frac{\exp\left(\tilde{V}^{E}/\sigma^{E}\right)}{\exp\left(\tilde{V}^{E}/\sigma^{E}\right) + \exp\left(\tilde{V}^{H}/\sigma^{E}\right)}$$

and

$$Pr\left(\ell_t^E\right) = \frac{\exp\left(\tilde{V}_t^{E,\ell_t^E}/(\sigma^{\ell^E}\lambda)\right)\left(\sum_{\ell\in\{0,PT,FT\}}\exp\left(\tilde{V}_t^{E,\ell}/(\sigma^{\ell^E}\lambda)\right)\right)^{\lambda-1}}{\left(\exp\left(\tilde{V}_t^{D}-\delta/\sigma^{\ell^E}\right)\right)+\left(\sum_{\ell\in\{0,PT,FT\}}\exp\left(\tilde{V}_t^{E,\ell}/(\sigma^{\ell^E}\lambda)\right)\right)^{\lambda}},$$

where σ^E and σ^{ℓ^E} are parameters governing the variance of the enrollment shock and college working shock, respectively, and λ is a nesting parameter and where value functions with tildes represent the value function minus the idiosyncratic preference draws.

The likelihood contribution of an agent who drops out in year T_i^D, has a sequence of work in college decisions of $\left\{\ell_{it}^E\right\}_{t=1}^{T^{dropout}-1}$, and has observed college transfers $tr_i^{E,o}$ is

$$\begin{aligned}\mathcal{L}_i\left(e_i = D, tr_i^{E,o}, \left\{\ell_{it}^E\right\}_{t=1}^{T_i^D-1}|\Gamma\right) = \\ Pr\left(E\right) f^N\left(\frac{tr_i^{E\star}-tr_i^{E,o}}{\sigma^{e^{tr}}}\right)\frac{1}{\sigma^{e^{tr}}}\left(\prod_{t=1}^{T_i^D-1} Pr\left(\ell_{it}^E\right)\right) Pr\left(D_{T^D}\right),\end{aligned} \tag{13}$$

where the probability of dropping out, $Pr\left(D_{T^D}\right)$, is given by the extreme value choice probabilities as

$$Pr\left(D_{T^D}\right) = \frac{\left(\exp\left(\tilde{V}_{T^D}^{D}-\delta/\sigma^{\ell^E}\right)\right)}{\left(\exp\left(\tilde{V}_{T^D}^{D}-\delta/\sigma^{\ell^E}\right)\right)+\left(\sum_{\ell\in\{0,PT,FT\}}\exp\left(\tilde{V}_{T^D}^{E,\ell}/(\sigma^{\ell^E}\lambda)\right)\right)^{\lambda}}.$$

The likelihood function of an agent who enters the labor force directly and is observed with transfers $tr_i^{H,o}$ is given by

$$\mathcal{L}_i\left(e_i = H, tr_i^{H,o}|\Gamma\right) = \left(1 - Pr\left(E\right)\right) f^N\left(\frac{tr_i^{H\star}-tr_i^{H,o}}{\sigma^{e^{tr}}}\right)\frac{1}{\sigma^{e^{tr}}}. \tag{14}$$

We therefore choose the parameters Γ to maximize the log likelihood:

$$\max_{\Gamma}\sum_i \log\mathcal{L}_i\left(\cdot|\Gamma\right).$$

3.5 Nonparametric Identification of Utility Function

In this section, we provide a formal discussion on how the psychic costs are nonparametrically identified. The argument closely follows Matzkin (1991, 1992, 1993) and Bajari et al. (2016).

Setup and Notation Consider an agent currently enrolled in college. The agent chooses between dropping out of college, and continuing in college and a labor supply quantity. Let $j = \{0, PT, FT, D\}$ index possible choices for an agent who is enrolled in college (enrolling and not working, enrolling and working part time, enrolling and working full time, and dropping out.) Agents possess a vector X of demographics and family income I. We write the choice specific value function associated with each continuation and work option ($j \in \{0, PT, FT\}$) as:

$$V_t^j \left(X, I, \tilde{a}_t, \varepsilon_t^j\right) = \max_{c_t} \left[\frac{c_t^{1-\gamma}}{1-\gamma} - h\left(X, j\right) + \varepsilon_t^j + \beta \mathbb{E}\left[V_{t+1}\left(\cdot\right)\right]\right]$$

subject to

$$c_t = \tilde{a}_t\left(j\right) - a_{t+1}$$

and the borrowing constraint. We define:

$$\tilde{a}_t\left(j\right) = \ell_t^E(j)\underline{\omega} + a_t\left(1 + r\left(a_t, I\right)\right) - \mathcal{F}(X) + \mathcal{G}\left(X, I\right) + tr_t^E(X, I, \mathcal{G}(X, I))$$

as the total assets available conditional on an agent's labor supply and let $\tilde{a}_t = \left(\tilde{a}_t\left(0\right), \tilde{a}_t\left(PT\right), \tilde{a}_t\left(FT\right), a_t\right) \in A$ give the vector of available assets associated with each choice. $h\left(X, j\right)$ is a continuous function in X that maps the agent's characteristics and choices into the deterministic component of psychic cost and ε_t^j are unobservable random terms representing the idiosyncratic component of psychic costs. Note that while the psychic cost function is allowed to depend on the vector of observable characteristics X, the vector of effective assets $\tilde{a}_t$ is excluded from the psychic cost function.

The choice specific value function for dropping out ($j = D$) is:

$$V_t^D(X, I, \tilde{a}_t, \varepsilon_t^D) = \mathbb{E}\left[V_t^W(X, I, e = D, a_t, w_t)\right] - d\left(X\right) + \varepsilon_t^D$$

where $d\left(X\right)$ is a continuous function which represents the deterministic component of the psychic cost of dropping out and ε_t^D is an unobservable random term representing the idiosyncratic component of the psychic cost of dropping out.

Let $\eta_t^j = \varepsilon_t^j - \varepsilon_t^D$ for $j \in \{0, PT, FT\}$ represent the idiosyncratic component of each continuation and labor supply option minus the idiosyncratic component of the psychic cost of dropping out. The distribution of $\eta_t = \left(\eta_t^0, \eta_t^{PT}, \eta_t^{FT}\right)$ is characterized by the joint CDF:

$$F_\eta\left(\tilde{\eta}\right) = P\left(\eta_t^0 \leq \tilde{\eta}^0, \eta_t^{PT} \leq \tilde{\eta}^{PT}, \eta_t^{FT} \leq \tilde{\eta}^{FT}\right).$$

We assume the distribution of η is uncorrelated with everything else in the model such that $F_\eta(\tilde{\eta}|X) = F_\eta(\tilde{\eta})$. Further, we assume that $F_\eta(\tilde{\eta})$ is continuous and strictly increasing. As is standard in discrete choice models, only the distribution of $\eta^j = \varepsilon^j - \varepsilon^D$ is identified, not the distribution of the ε's themselves. Further, let $\hat{h}(X,j) = h(X,j) - d(X)$ be the difference between the psychic cost function associated with continuation in college and the psychic cost function associated with dropping out. We assume the discount factor, β, is known.

Identification In the quantitative version of the model, we made parametric assumptions about the functions $h(X,j)$ and $d(X)$ and the distribution of the idiosyncratic preference shocks. Here we show the conditions under which the model is identified without these parametric and distributional assumptions.

We prove identification in four steps. First, we focus on agents in their final year of college. Let this year be denoted by $t = T_G$. We show that the parameter γ and the distribution of η can be identified from variation in choice probabilities and effective assets of agents in their final year of college ($t = T_G$), holding demographics X and parental income I constant. Second, once F_η is identified, we can use variation in X to identify $\hat{h}(X,j)$, the differences in the common component of psychic costs. Third, once F_η and $\hat{h}(X,j)$ are identified, we show that $d(X)$ is identified by the dropout probabilities of agents in the penultimate year of college ($t = T_G - 1$). Finally, we show that the distribution of the initial enrollment shock is identified by the probabilities of college enrollment.

Consider individuals in their sixth year of college who graduate with certainty at the end of the period. Note that we can rewrite the choice specific Bellman equation for each work in college option ($j \in \{0, PT, FT\}$) in year $t = T_G$ as

$$V_{T_G}^j\left(X, I, \tilde{a}_{T_G}, \varepsilon_{T_G}^j\right) = \tilde{V}_{T_G}^j\left(X, I, \tilde{a}_{T_G} : \gamma\right) - h(X,j) + \varepsilon_{T_G}^j$$

where we define

$$\tilde{V}_{T_G}^j\left(X, I, \tilde{a}_{T_G} : \gamma\right) = \max_{c_{T_G}}\left[\frac{c_{T_G}^{1-\gamma}}{1-\gamma} + \beta\mathbb{E}\left[V_{T_G+1}^W\left(X, I, e = G, a_{T_G+1}, w_{T_G+1}\right)\right]\right]$$

subject to the budget constraint and borrowing constraint. Since the parameters of the wage equation are identified directly from the earnings data,[15] the function $\tilde{V}_{T_G}^j(\cdot : \gamma)$ is known up to the parameter γ. Similarly, we can write the value function associated with dropping out as:

$$V_{T_G}^D(X, I, \tilde{a}_{T_G}, \varepsilon_{T_G}^D) = \tilde{V}_{T_G}^D\left(X, I, \tilde{a}_{T_G} : \gamma\right) - d(X) + \varepsilon_{T_G}^D$$

where

$$\tilde{V}_{T_G}^D\left(X, I, \tilde{a}_{T_G} : \gamma\right) = \mathbb{E}\left[V_{T_G}^W\left(X, I, e = D, a_{T_G}, w_{T_G}\right)\right]$$

[15]This is because the wage equation is only a function of observables and idiosyncratic shocks that are uncorrelated with everything else in the model.

is a known function up to the parameter γ.[16]

The relationship between the choice probabilities and the vector $\tilde{a}_{T_G}$, holding X and I constant identifies γ. To see this, note that the probability of dropping out of college is given by

$$P\left(D|X,I,\tilde{a}_{T_G}\right) = \\ \text{Prob}\Big(\tilde{V}^{D}_{T_G}\left(X,I,\tilde{a}_{T_G}:\gamma\right) - d\left(X\right) - \left(\tilde{V}^{j'}_{T_G}\left(X,I,\tilde{a}_{T_G}\left(j'\right):\gamma\right) - h\left(X,j'\right)\right) \geq \eta^{j'}_{T_G} \\ \forall j' \in \{0,PT,FT\}\Big)$$

and the probability of choosing $j = \{0, PT, FT\}$ is given by

$$P\left(j|X,I,\tilde{a}_{T_G}\right) = \\ \text{Prob}\Big(\tilde{V}^{D}_{T_G}\left(X,I,\tilde{a}_{T_G}:\gamma\right) - d\left(X\right) - \left(\tilde{V}^{j}_{T_G}\left(X,I,\tilde{a}_{T_G}\left(j\right):\gamma\right) - h\left(X,j\right)\right) \leq \eta^{j}_{T_G}, \\ \tilde{V}^{j}_{T_G}\left(X,I,\tilde{a}_{T_G}\left(j\right):\gamma\right) - h\left(X,j\right) - \left(\tilde{V}^{j'}_{T_G}\left(X,I,\tilde{a}_{T_G}\left(j'\right):\gamma\right) - h\left(X,j'\right)\right) \geq \eta^{j'}_{T_G} - \eta^{j}_{T_G} \\ \forall j' \in \{0,PT,FT\} \setminus \{j\}\Big)$$

Therefore, any vector $\tilde{a}_{T_G}$ such that

$$\tilde{V}^{D}_{T_G}\left(X,I,\tilde{a}_{T_G}:\gamma\right) - \tilde{V}^{j}_{T_G}\left(X,I,\tilde{a}_{T_G}\left(j\right):\gamma\right) = k_j \;\; \forall \{0,PT,FT\}$$

where $k = (k_0, k_{PT}, k_{FT})$ is vector of constants implies the same set of choice probabilities, holding X and I constant. Consider two vectors $\tilde{a}'_{T_g}$ and $\tilde{a}''_{T_g}$ such that $P\left(j|X,I,\tilde{a}'_{T_G}\right) = P\left(j|X,I,\tilde{a}''_{T_g}\right)$ for each $j \in \{0, PT, FT, D\}$. Then we must have:

$$\tilde{V}^{D}_{T_G}\left(X,I,\tilde{a}'_{T_G}:\gamma\right) - \tilde{V}^{j}_{T_G}\left(X,I,\tilde{a}'_{T_G}\left(j\right):\gamma\right) = \tilde{V}^{D}_{T_G}\left(X,I,\tilde{a}''_{T_G}:\gamma\right) - \tilde{V}^{j}_{T_G}\left(X,I,\tilde{a}''_{T_G}\left(j\right):\gamma\right) \\ \forall j \in \{0,PT,FT\}$$

This identifies γ.

Once γ is identified, the deterministic portions of the utility functions, $(\tilde{V}^{D}_{T_G}\left(X,I,\tilde{a}_{T_G}\right) - d\left(X\right))$ and $\tilde{V}^{j}_{T_G}\left(X,I,\tilde{a}_{T_G}\right) - h\left(X,j\right)$ are additively separable into a function that is known and a continuous function of the remaining variables. We remove the dependence of $\tilde{V}^{D}_{T_G}\left(\cdot\right)$ and $\tilde{V}^{j}_{T_G}\left(\cdot\right)$ on γ now that γ is identified.

We now turn to the identification of the psychic costs. Without any further normalizations, the mean of the idiosyncratic components of psychic costs are not separately identified from additive constants in the psychic cost function h and the dropout function d. We therefore need

[16]This is true because all of the parameters which determine wages and the disutility of labor supply in the labor market are known.

to make four normalizations, one for each option $j \in \{0, PT, FT, D\}$. First, we normalize the idiosyncratic shock associated with dropping out of college to be mean 0: $\mathbb{E}\left(\varepsilon^D\right) = 0$. Next, let $\bar{X}$ denote a particular value of X that is known to the econometrician. We normalize $\hat{h}\left(\bar{X}, j\right) = \mu_j$, where μ_j is a known value for all $j \in \{0, PT, FT\}$.[17]

Let

$$\begin{aligned}
&\hat{V}^D_{T_G}(X, I, \tilde{a}_{T_G}) = \\
&\left(\tilde{V}^D_{T_G}(X, I, \tilde{a}_{T_G}) - \left(\tilde{V}^0_{T_G}(X, I, \tilde{a}_{T_G}) - \hat{h}(X, j=0)\right),\right. \\
&\tilde{V}^D_{T_G}(X, I, \tilde{a}_{T_G}) - \left(\tilde{V}^{PT}_{T_G}(X, I, \tilde{a}_{T_G}) - \hat{h}(X, j=PT)\right), \\
&\left.\tilde{V}^D_{T_G}(X, I, \tilde{a}_{T_G}) - \left(\tilde{V}^{FT}_{T_G}(X, I, \tilde{a}_{T_G}) - \hat{h}(X, j=FT)\right)\right).
\end{aligned}$$

be the 3 dimensional vector which gives the deterministic portion of utility associated with dropping out minus the deterministic utility of each of the other three options. Fixing X at the known vector $\bar{X}$, we can write these objects as:

$$\begin{aligned}
&\hat{V}^D_{T_G}\left(\bar{X}, I, \tilde{a}_{T_G}\right) = \\
&\left(\tilde{V}^D_{T_G}\left(\bar{X}, I, \tilde{a}_{T_G}\right) - \left(\tilde{V}^0_{T_G}\left(\bar{X}, I, \tilde{a}_{T_G}\right) - \mu_0\right),\right. \\
&\tilde{V}^D_{T_G}\left(\bar{X}, I, \tilde{a}_{T_G}\right) - \left(\tilde{V}^{PT}_{T_G}(X, I, \tilde{a}_{T_G}) - \mu_{PT}\right), \\
&\left.\tilde{V}^D_{T_G}\left(\bar{X}, I, \tilde{a}_{T_G}\right) - \left(\tilde{V}^{FT}_{T_G}\left(\bar{X}, I, \tilde{a}_{T_G}\right) - \mu_{FT}\right)\right).
\end{aligned}$$

The vector $\hat{V}^D_{T_G}\left(\bar{X}, I, \tilde{a}_{T_G}\right)$ is known by the econometrician, given that the function $\tilde{V}^j_{T_G}(\cdot)$ is known for all j and because μ_j is a known value for all j.

The identification of $\hat{h}$ and F_η then follows from the arguments in Matzkin (1993).[18] The probability of dropping out can be written as:

$$P\left(D|\bar{X}, I, \tilde{a}_{T_g}\right) = F_\eta\left(\hat{V}^D_{T_G}\left(\bar{X}, I, \tilde{a}_{T_g}\right)\right).$$

[17] In the quantitative version of the model, we do not normalize the level of the deterministic portion of psychic cost, which is dictated by the parameter κ_0. Instead, we assume that the idiosyncratic component of psychic costs are distributed as nested Logit, which normalizes the mean of the idiosyncratic components of psychic costs. We make a different normalization here by normalizing the level of the deterministic portion of the psychic cost, as it makes the proof a bit easier to follow.

[18] See Theorem 2.

Assume for all $\tilde{\eta} \in R^3$, there exists a vector $\tilde{a}_{T_G}$ with positive density conditional on $\bar{X}$ such that $\hat{V}^D_{T_G}\left(\bar{X}, I, \tilde{a}_{T_G}\right) = \tilde{\eta}$.[19] Then F_η can be recovered from choice probabilities.

Once the distribution F_η is known, we can recover the function $\hat{h}$ as in Matzkin (1991): the choice probabilities at different levels of X identify the respective values of $\hat{h}$.

This leaves us with identification of $d(X)$ and the distribution of the initial enrollment cost ε^E. First, we consider the identification of the dropout cost $d(X)$. Note that we have already identified the function $\hat{h}$, which determines the deterministic portion of utility differences across choices of enrolled agents. The function $d(X)$ pins down the *level* of utility of enrolled agents. Specifically, holding $\hat{h}$ constant, increasing $d(X)$ lowers the flow utility levels of all choices for enrolled agents by the same amount. In our model, dropping out plays the role of a *terminating action*—an action which ends the agent's dynamic discrete choice problem. Therefore, the level of utility of enrolled agents, which is determined by the function $d(X)$, is identified following the arguments of Bajari et al. (2016). In essence, if an agent chooses to drop out in the current period, she will not be enrolled in college in the following period. Therefore, a higher level of utility associated with being enrolled in college increases the value of not dropping out and therefore decreases the likelihood of dropping out.

To see this explicitly, consider the problem of an agent in the penultimate year of college, $T_G - 1$. For ease of exposition, we assume that $Pr^{Grad}_{T_G-1} = 0$ but the argument is the same with positive probability of graduation. We can write the agent's choice specific utility functions in the year $T_G - 1$ as

$$V^j_{T_G-1}\left(X, I, \tilde{a}_{T_G-1}, \varepsilon^j_{T_G-1}\right) = \max_{c_t}\left[\frac{c_t^{1-\gamma}}{1-\gamma} - \hat{h}\left(X, j\right) + \varepsilon^j_{T_G-1} - d\left(X\right) + \beta\mathbb{E}\left[V^E_{T_G}\left(X, I, a_{T_G}, \varepsilon_{T_G}\right)\right]\right]$$

for $j \in \{0, PT, FT\}$. The continuation value is given by

$$V^E_{T_G}(X, I, a_{T_G}, \varepsilon_{T_G}) = \max_{j\in\{0,PT,FT,D\}}[V^j_{T_G}(X, I, a_{T_G}, \varepsilon_{T_G})].$$

The choice specific utility functions in period T_G can be written as

$$V^j_{T_G}\left(X, I, \tilde{a}_{T_G}, \varepsilon^j_{T_G}\right) = \tilde{V}^j_{T_G}\left(X, I, \tilde{a}_{T_G}\right) - \hat{h}\left(X, j\right) + \left(\eta^j_{T_G} + \varepsilon^D_{T_G}\right) - d(X)$$

for $j \in \{0, PT, FT\}$ and

$$V^D_{T_G}(X, I, \tilde{a}_{T_G}, \varepsilon^D_{T_G}) = \tilde{V}^D_{T_G}\left(X, I, \tilde{a}_{T_G}\right) + \varepsilon^D_{T_G} - d\left(X\right).$$

[19] We do not expect this assumption to hold in practice, given that we have finite data and there is limited variation in assets in the model. As such, we make distributional and parametric assumption on the psychic costs in our quantitative model. The goal of this exercise is to state the necessary conditions for non-parametric identification.

Note that $-d(X)$, which pins down the level of utility of an agent enrolled in college, enters additively into $V^j_{T_G}(\cdot)$, for each option j. We can therefore write the expectation of continuation value as

$$\mathbb{E}\left[V^E_{T_G}(X, I, a_{T_G}, \varepsilon_{T_G})\right] = \tilde{EV}^E_{T_G}(X, I, a_{T_G}, \varepsilon_{T_G}) - d(X)$$

where we define

$$\begin{aligned}
&\widetilde{EV}^E_{T_G}(X, I, a_{T_G}) = \mathbb{E}\Bigg[\max\Bigg\{\\
&\tilde{V}^0_{T_G}(X, I, \tilde{a}_{T_G}) - \hat{h}(X, j=0) + \eta^0_{T_G}, \tilde{V}^{PT}_{T_G}(X, I, \tilde{a}_{T_G}) - \hat{h}(X, j=PT) + \eta^{PT}_{T_G},\\
&\tilde{V}^{FT}_{T_G}(X, I, \tilde{a}_{T_G}) - \hat{h}(X, j=FT) + \eta^{FT}_{T_G}, \tilde{V}^D_{T_G}(X, I, \tilde{a}_{T_G})\Bigg\}\Bigg]
\end{aligned}$$

as the expected value of being enrolled in college next year less the term $-d(X)$, where we have used the normalization that $\mathbb{E}\left(\varepsilon^D\right) = 0$. Note that the function $\widetilde{EV}^E_{T_G}(X, I, a_t)$ is fully known by the econometrician, given that 1) the function $\tilde{V}^j_{T_G}(\cdot)$ is known for all j, 2) the function $\hat{h}(\cdot)$ is known, and 3) the distribution F_η is known. We can then rewrite the choice specific value functions for $j \in \{0, PT, FT\}$ in the penultimate period as

$$V^j_{T_G-1}\left(X, I, \tilde{a}_{T_G-1}, \varepsilon^j_{T_G-1}\right) = \tilde{V}^j_{T_G-1}(X, I, \tilde{a}_{T_G-1}) - \hat{h}(X, j) + \varepsilon^j_{T_G-1} - d(X) - \beta d(X),$$

where the function we define as

$$\tilde{V}^j_{T_G-1}(X, I, \tilde{a}_{T_G-1}) = \max_{c_t}\left[\frac{c_t^{1-\gamma}}{1-\gamma} + \beta\widetilde{EV}^E_{T_G}(X, I, a_{T_G})\right]$$

is fully known to the econometrician.

Further, in the penultimate period of college, the value of dropping out is given by

$$V^D_{T_G-1}(X, I, \tilde{a}_{T_G-1}, \varepsilon^D_{T_G-1}) = \tilde{V}^D_{T_G-1}(X, I, \tilde{a}_{T_G-1}) + \varepsilon^D_{T_G-1} - d(X).$$

where we $\tilde{V}^D_{T_G-1}(X, I, \tilde{a}_t) = \mathbb{E}\left[V^W_{T_G-1}(X, I, e=D, a_{T_G-1}, w_{T_G-1})\right]$ is a known function. The difference in the value of dropping out and the value of $j \in \{0, PT, FT\}$ is given by

$$V^D_{T_G-1}(\cdot) - V^j_{T_G-1}(\cdot) = \tilde{V}^D_{T_G-1}(X, I, \tilde{a}_{T_g-1}) - \left(\tilde{V}^j_{T_G-1}(X, I, \tilde{a}_{T_G-1}) - \hat{h}(X, j) + \eta^j_{T_G-1} - \beta d(X)\right).$$

The probability of dropping out can then be written as:

$$
\begin{aligned}
&P\left(D|X,I,\tilde{a}_{T_G-1}\right)= \\
&F_\eta\Bigg(\tilde{V}^{D}_{T_G-1}\left(X,I,\tilde{a}_{T_G-1}\right)-\left(\tilde{V}^{0}_{T_G-1}\left(X,I,\tilde{a}_{T_G-1}\right)-\hat{h}\left(X,j=0\right)-\beta d\left(X\right)\right)\Big), \\
&\tilde{V}^{D}_{T_G-1}\left(X,I,\tilde{a}_{T_G-1}\right)-\left(\tilde{V}^{PT}_{T_G-1}\left(X,I,\tilde{a}_{T_G-1}\right)-\hat{h}\left(X,j=PT\right)-\beta d\left(X\right)\right), \\
&\tilde{V}^{D}_{T_G-1}\left(X,I,\tilde{a}_{T_G-1}\right)-\left(\tilde{V}^{FT}_{T_G-1}\left(X,I,\tilde{a}_{T_G-1}\right)-\hat{h}\left(X,j=FT\right)-\beta d\left(X\right)\right)\Bigg).
\end{aligned}
$$

Therefore, we can again use the arguments in Matzkin (1991) and Matzkin (1993) to back out the function $d(X)$ given that the joint distribution of η, the discount factor β, and the $\hat{h}(X,j)$ and $\tilde{V}^{j}_{T_G-1}$ functions are all known.

Finally, we turn to the initial enrollment cost. The agent's initial enrollment decision can be written as the binary threshold crossing model:

$$
\mathbb{E}\left[V^{E}_{1}\left(X,I,a_1=0,\varepsilon_1\right)\right]-\mathbb{E}\left[V^{W}_{1}(X,I,e=H,a_1=0,w_1)\right]+\varepsilon^{E}>0,
$$

where $\left[V^{E}_{1}(\cdot)\right]$ is the expected value of a first year college enrollee and $\mathbb{E}\left[V^{W}_{1}(\cdot)\right]$ is the expected value of directly entering the labor market. The functions $\mathbb{E}\left[V^{E}_{1}(\cdot)\right]$ and $\mathbb{E}\left[V^{W}_{1}(\cdot)\right]$ are fully known and continuous. We can therefore identify the distribution of the psychic enrollment cost using the arguments in Matzkin (1992).

4 Additional Graphs on Model Fit

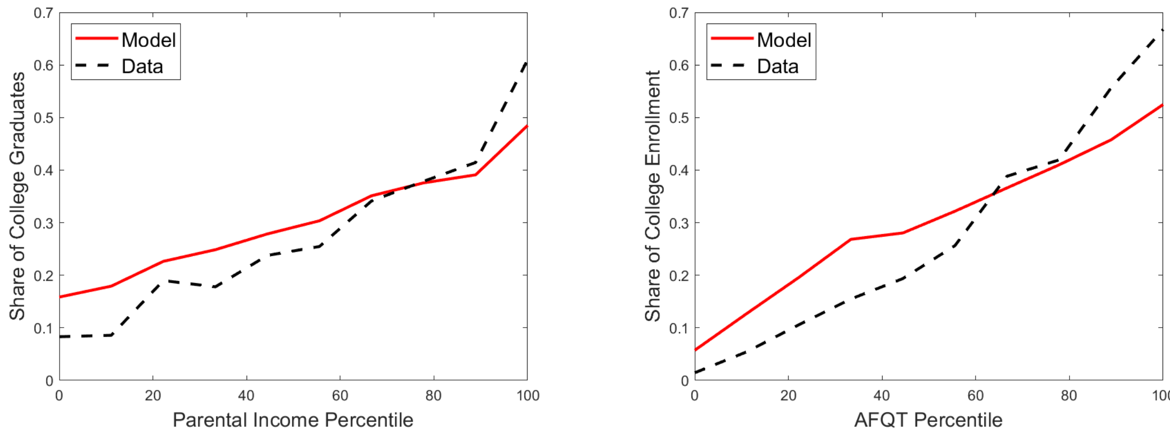
4.1 Graduation Rates and Enrollment by Gender

Figure 1 shows the college graduation rates as a function of parental income and ability in the model and in the data. The model is able to replicate these moments well.

Figure 2(a) shows the college enrollment rates for male and female students as a function of parental income in the model and in the data. Figure 2(b) shows the college enrollment rates for male and female students as a function of ability in the model and in the data. We can see that the model is able to replicate these moments quite well.

4.2 Graduation and Dropout Over Time

Figure 3 shows graduation and dropout fractions over time in the model and the data. The solid red line and the dashed black line show the fraction of the total population that have graduated as a function of number of years of college completed in the model and the data, respectively. In



(a) Graduation Rates and Parental Income

(b) Graduation Rates and AFQT

Figure 1: Graduation and Enrollment Rates

Notes: The solid (red) line shows simulated enrollment shares by parental income and AFQT percentile. This is compared to the dashed (black) line which shows the shares in the data.

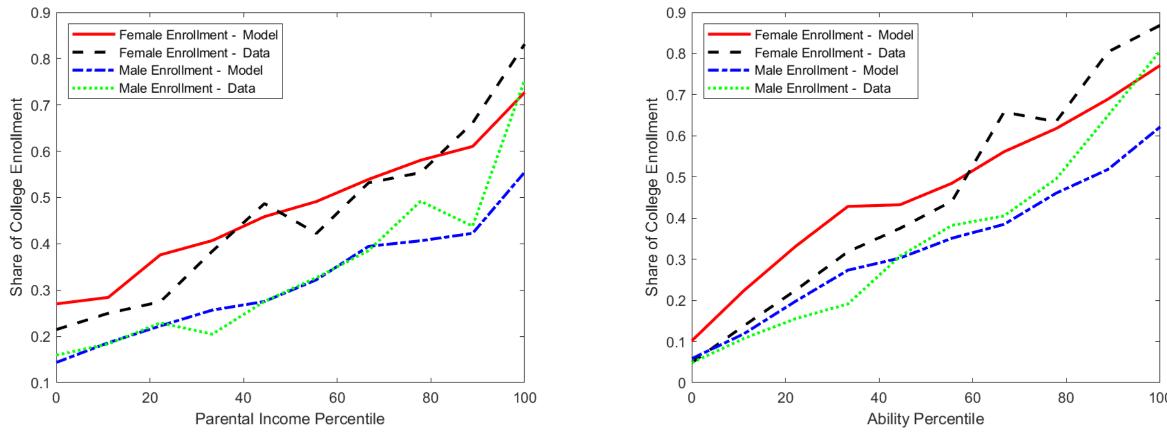
both the model and the data, graduation rates are very low for students with less than three years of college. Graduation shares peak at four years before decreasing. The dashed-dotted blue line and the dotted green line show the fraction of students that drop out in each year in the model and data, respectively. Dropout shares are slightly downward sloping as a function of years in college in both the model and the data. This slope is slightly steeper in the model compared to the data.

4.3 Parental Transfers

We analyze the fit of our model with respect to parental transfers in Figure 4. We can see that college transfers are strongly increasing in parental income in both the model and data, though our model slightly underestimates the average college transfers in the data.

4.4 Earnings and College Premia

Table 3 analyzes the performance of the model with respect to earnings dynamics. We can only compare the model to the NLSY97 data up to age 34 since cohorts in the NLSY97 are born between 1980 and 1984. The simulated mean earnings across ages are very close to those in the data. As described in Section 4, we account for top-coding of earnings data by appending Pareto tails to the observed earnings distribution. As such, average earnings are slightly larger in model as compared to the data. We match college earnings premia very closely until around



(a) Enrollment Rates and Parental Income

(b) Enrollment Rates and AFQT

Figure 2: Graduation and Enrollment Rates by Gender

Notes: The panel on the left shows the relationship between enrollment rates and parental income in the model and in the data for females and males. The panel on the right shows the relationship between enrollment rates and ability in the model and in the data for females and males.

Age	**Mean Earnings** *High-School*		*College*		**College Premia**		**SD(log (y))**	
j	Model	Data	Model	Data	Model	Data	Model	Data
25	22,938	21,348	26,923	25,205	1.17	1.18	0.66	0.59
26	23,747	22,407	29,353	28,300	1.24	1.26	0.67	0.60
27	24,549	23,340	31,829	31,781	1.30	1.36	0.67	0.61
28	25,340	24,022	34,334	33,840	1.35	1.41	0.68	0.62
29	26,117	25,217	36,848	36,254	1.41	1.44	0.69	0.65
30	26,877	25,306	39,354	37,904	1.46	1.50	0.70	0.65
31	27,617	26,449	41,833	40,904	1.51	1.55	0.70	0.66
32	28,334	27,346	44,267	42,954	1.56	1.57	0.71	0.67
33	29,025	28,680	46,639	44,346	1.61	1.55	0.72	0.68
34	29,687	30,494	48,932	46,872	1.65	1.54	0.72	0.67

Notes: Data based on NLSY97 with cohorts born between 1980 and 1984. Mean earnings expressed in year 2000 dollars. Most recent wave from 2015. Model based moment results represent results from estimated model. Zero and small earnings below $300 a month excluded. SD(log y) equal to standard deviation of log earnings. NLSY97 is top coded at income levels around $155,000.

Table 3: Earnings Dynamics

age 32. After that, the model and data diverge slightly as more and more college students reach top-coded earnings in the NLSY97.[20]

[20]The effect of the fatter right tails we include in the model can also be seen in the fit of standard deviation of log earnings. The simulated standard deviation of log earnings is 4-7 log points higher than that in the data from age 25 to age 34.

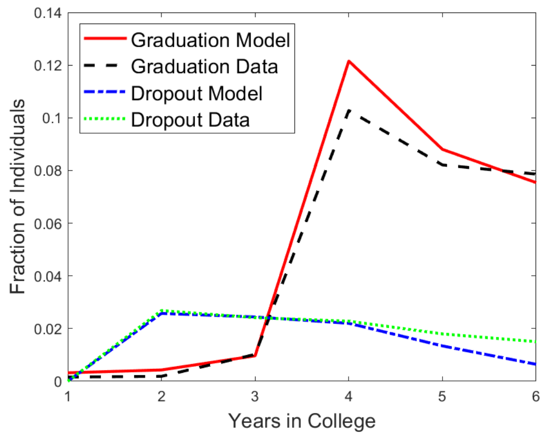


Figure 3: Model Fit: Graduation and Dropout Over Time

Notes: The figure shows simulated graduation and dropout rates in the model versus the NLSY97.

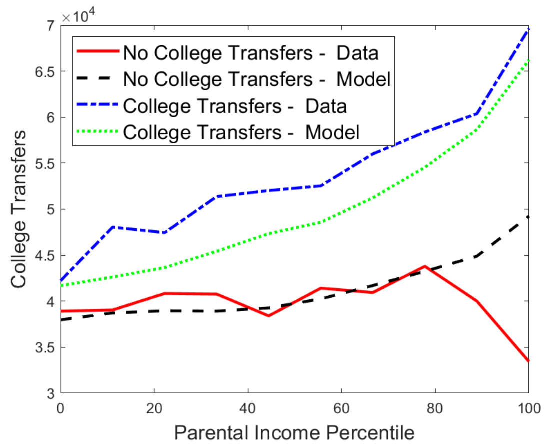


Figure 4: Model Fit: College Transfers and Parental Income

Notes: This figure on the right shows the present value of parental transfers given by parents of college enrollees and non-enrollees in data (NLSY97) versus model.

4.5 Earnings Profiles Model

Figure 5 shows the simulated average for college graduates and high school graduates as a function of age.

4.6 Untargeted Moments

Responsiveness of Enrollment to Grant Increases. Many papers have analyzed the impact of increases in grants or decreases in tuition on college enrollment. Deming and Dynarski (2009) survey the literature. The estimated impact of a \$1,000 increase in yearly grants (or a respective reduction in tuition) on enrollment ranges from 1 to 6 percentage points, depending on the policy

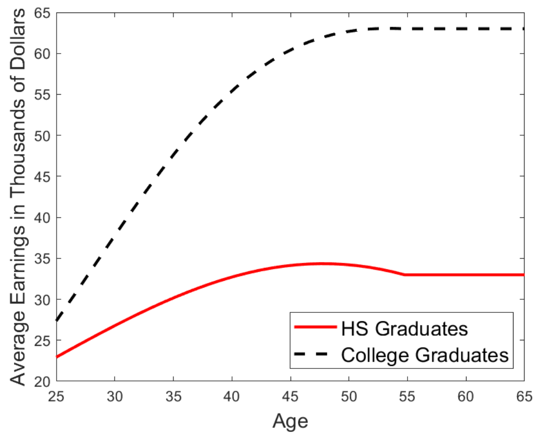


Figure 5: College and High School Graduate Earnings Profiles.

reform and research design. A more recent study by Castleman and Long (2016) 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.

Simulating a $1,000 increase in financial aid for all individuals in our model leads to a 1.69 percentage point increase in overall enrollment rates and a 2.06 percentage point increase for students near the studied discontinuity in Castleman and Long (2016). Overall, our simulated elasticities are fairly consistent with these reduced-form estimates. This gives us confidence in our maximum likelihood estimates, especially given that these reduced form estimates were not targeted in estimation.

Importance of Parental Income. Individuals with higher parental income are more likely to receive a college degree. However, it is not obvious whether this is primarily driven by parental income itself or by variables correlated with parental income and college graduation. Using income tax data and a research design exploiting parental layoffs, Hilger (2016) finds that a $1,000 increase in parental income leads to an increase in college enrollment of 0.43 percentage points. To test our model, we increased parental income for each individual by $1,000 and obtained increases in college enrollment by 0.18 percentage points. Our model predicts a moderate direct effect of parental income, smaller but in line with Hilger (2016).

Returns for Marginal Students. We find a return to one year of schooling of 12.1% for marginal students. This reflects that marginal students are of lower ability on average than inframarginal students and is also in line with Oreopoulos and Petronijevic (2013). A clean way to infer returns for marginal students is found in Zimmerman (2014). In his study, students are marginal with respect to academic ability, measured by a GPA admission cutoff. He finds that these students have earnings 22% higher than those just below the cutoff, when earnings are measured 8 to 14 years after high school graduation. We perform a similar simulation and make use of the fact that the NLSY also provides GPA data. In fact, our model gives a return to college of 26.3%, measured 8 to 14 years after high school graduation, for students with a GPA in this neighborhood.[21]

5 Additional Decompositions

5.1 Marginal and Inframarginals Evaluated at Current Financial Aid Levels

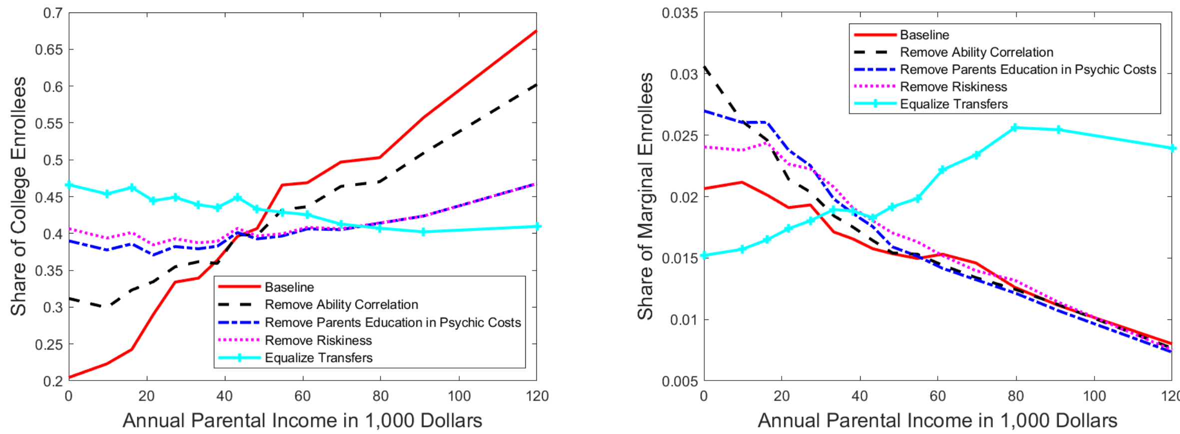
In the main text, we plotted the share of marginal enrollees and inframarginal enrollees at a flat financial schedule for a number of model specifications. In this section, we repeat this exercise but plot the share of marginal enrollees and inframarginal enrollees at the current financial aid schedule. The results are very similar. The relationship between parental income and the share of inframarginal students has become weaker (and eventually becomes negative), reflecting that the current financial aid schedule is decreasing in parental income, see Figure 6(a). Further, children with high income parents are more likely to be marginal with respect to financial aid relative to graph in the main text, again reflecting that they receive less financial aid than children with low income parents, see Figure 6(b).

5.2 An Alternative Decomposition: Different Order

In the main paper, we perform a model-based decomposition exercise to better understand which drive the optimal progressivity result. In this appendix, we perform a similar decomposition but alter the order in which we change various components to the model. In particular, we first remove the relationship between parental income and parental transfers, before proceeding to remove the relation between parental income and ability and the relation between parental education and the psychic costs of college. As before, all changes to the model specification are cumulative.

We first analyze the determinants of the positive relation between college enrollment and parental income in Figure 7(a) and the negative relationship between share of marginal students and parental income in Figure 7(a). The simulated relationships at a flat financial aid schedule are shown in the solid lines in the two figures. In this baseline case, college enrollment rates are strongly increasing in parental income while the share of marginal students are strongly decreasing in parental income. Next, in the turquoise lines, we set parental transfers exogenously to the mean

[21] 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 (2016) for disability insurance). Further, the fiscal costs of Medicare are likely to be much lower for individuals with a 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, 2016). These facts would generally strengthen the case for an increase in college subsidies.



(a) Inframarginal Enrollees

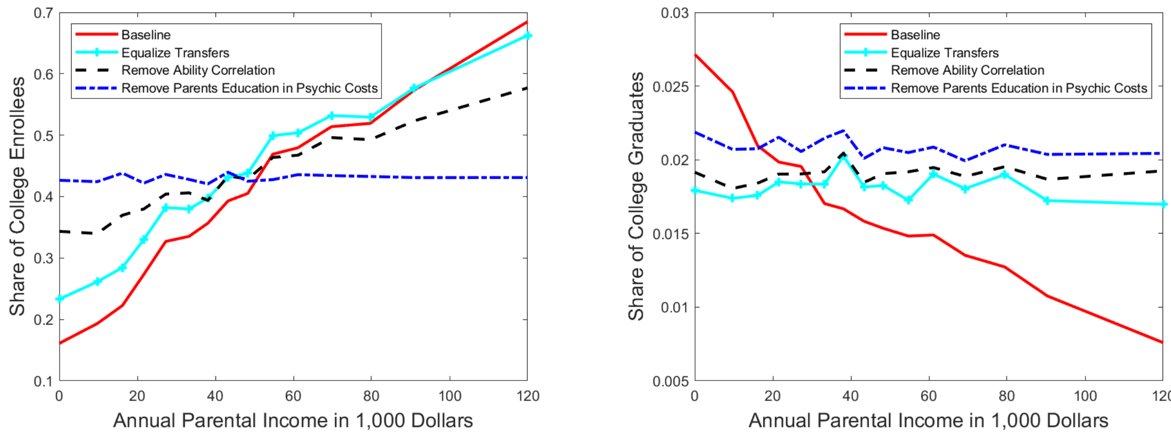
(b) Marginal Students

Figure 6: Model-Based Decomposition for Marginal and Inframarginal Students at Current Grant Schedule

Notes: We plot the share of college enrollees and marginal college enrollees given the current US aid schedule for different model specifications. The solid red line represent the baseline model. For the dashed black line we simulate a model version for which we remove the correlation between ability and parental income. For the dashed-dotted blue line we simulate a model version for which we additionally remove the correlation between the psychic costs and parental education. For the dotted pink line we simulate a model version for which additionally removes labor market riskiness; i.e. education decisions are made with no uncertainty about future wages. For the turquoise line with crosses we simulate a model version for which we set parental transfers to the mean parental transfers in the data, conditional on education.

levels for enrollees and non-enrollees and assume no families are eligible for subsidized Stafford loans. From Figure 7(a) we can see that the positive relation between college enrollment and parental income weakens slightly. The relation between parental income and share of marginal enrollees, however, flattens completely. The black dotted line and the blue dash-dotted line show the cases in which we remove the correlation between parental income and ability and in which we remove the relation between parental education and psychic costs, respectively. After removing these two factors there is no interesting heterogeneity between parental income groups. Removing these two relationships both weak the relationship between parental income and enrollment. In both these simulations, the gradient between parental income and share of marginal students remains flat.

We now simulate the respective optimal financial aid schedule under each model specification in Figure 8. When we remove the relationship between parental income and parental transfers (the turquoise line in Figure 8), the optimal financial aid schedule flattens. This flattening of the optimal schedule occurs because the relationship between parental income and share of marginal enrollees is flat. However, there optimal aid is still positive—ranging from above \$6,500 for the



(a) Inframarginal Graduates (b) Marginal Students

Figure 7: Alternative Model-Based Decomposition for Marginal and Inframarginal Students at Flat Grant Schedule

Notes: We plot the share of college enrollees and marginal college enrollees given a flat financial aid schedule for different model specifications. The solid red line represent the full model. For the turquoise line with crosses we simulate a model version for which set parental transfers to the mean parental transfers in the data, conditional on education. For the dashed black line we simulate a model version for which we remove the correlation between ability and parental income. For the dashed-dotted blue line we simulate a model version for which we additionally remove the correlation between the psychic costs and parental education.

poorest families to below \$4,500 for the wealtheat families. The optimal aid schedule is progressive because high income children are still much more likely inframarginal. When we remove the positive ability-income correlation (the black dashed line) and the relationship between parental education and psychic costs we flatten the relationship between parental income and share of inframarginal students. The optimal aid schedule flattens as a result.

5.3 Decomposition with Removal of Borrowing Constraints

In Figures 9(a) and 9(b), we perform the same decomposition as in the main paper but additionally remove borrowing constraints before equalizing parental transfers. We additionally assume that no families are eligible for subsidized Stafford loans throughout the decomposition. Figure 10 shows the resulting optimal financial aid under each model specification.

As before, the red line shows the baseline case, the black dotted line shows the case where we remove the ability correlation, the blue dash-dotted line shows the case where we remove the correlation between psychic cost of parental education, and the dotted pink line show the case with no labor market uncertainty. These lines tell essentially the same story as the decomposition in the main body. The green dotted lines show the case in which we remove borrowing constraints. As a result, the number of inframarginal students increases for all income groups, as student no

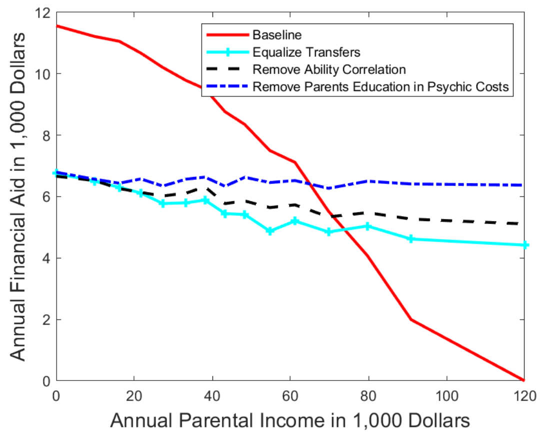


Figure 8: Optimal Financial Aid for Different Model Specifications

Notes: For each model specification (see Figure 7), we illustrate the respective optimal financial aid schedule.

longer have to deal with borrowing constraints in college. Additionally, the share of marginal enrollees drops substantially for all parental income groups. As students are no longer affected by borrowing constraints in college, the marginal benefit of additional financial aid decreases substantially.

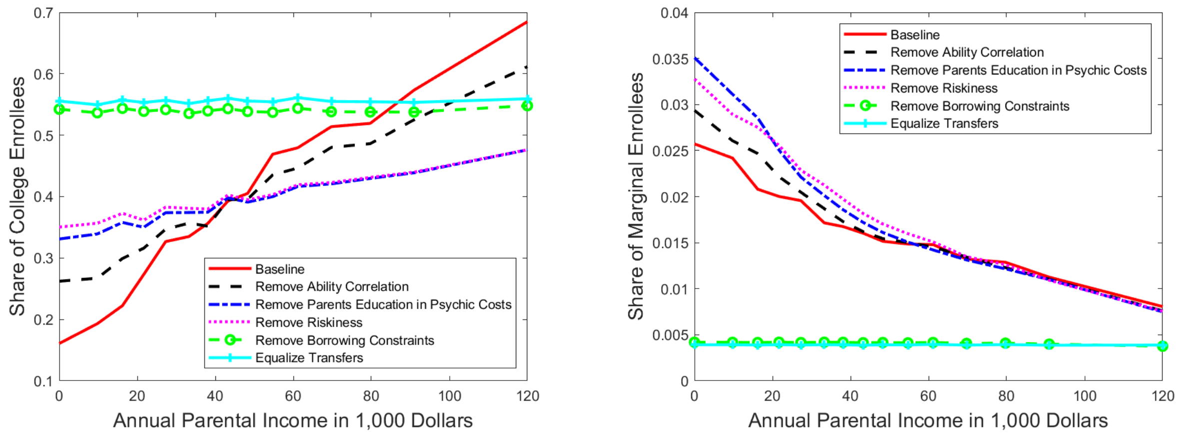
However, despite the fact that both the gradients of marginal and inframarginal enrollees are flat, the optimal aid is still slightly decreasing in parental income. This is because, once the correlations of parental income with marginal and inframarginal students have been shut down, the differences in marginal social welfare weights play a role. We find that at the flat financial aid schedule, the marginal social welfare weight of the poorest children is roughly 20% higher than that of the richest students. Essentially, given that enrollment is so unresponsive to financial aid, the social planner allocates financial aid to agents with the highest marginal social welfare weights. This leads to a slightly progressive financial aid schedule.

Equalizing parental transfers on top of this removes these differences in marginal social welfare weights and therefore leads to an flat optimal aid schedule.

6 Robustness and Additional Results

6.1 The Role of Borrowing Constraints

Figure 11(a) shows the optimal financial aid policies when we have abolished borrowing constraints. We first remove borrowing constraints and keep the current financial aid system. This will increase college enrollment and imply a windfall fiscal gain for the government. In a second



(a) Inframarginal Graduates (b) Marginal Students

Figure 9: Model-Based Decomposition for Marginal and Inframarginal Students at Flat Grant Schedule with Removal of Borrowing Constraints

Notes: We plot the share of college enrollees and marginal college enrollees given a flat financial aid schedule for different model specifications. We assume no subsidized Stafford loans for all specifications. The solid red line represent the full model at the flat financial aid schedule. For the dashed black line we simulate a model version for which we remove the correlation between ability and parental income. For the dashed-dotted blue line we simulate a model version for which we additionally remove the correlation between the psychic costs and parental education. For the dotted pink line we simulate a model version for which on top removes any riskiness; i.e. education decisions are made under perfect foresight. For the dashed green line with circles we simulate a model version for which on top we remove all borrowing constraints. For the turquoise line with crosses we simulate a model version for which set parental transfers to the mean parental transfers in the data, conditional on education.

step, we choose optimal financial aid but restrict the government to not use this windfall gain. Figure 11(b), shows the implied graduate patterns.

6.2 Varying Borrowing Constraints

To get a sense of how varying borrowing constraints would affect our main conclusions, we have re-estimated a version of the model in which the borrowing limit depends on parental resources. Here, it was very hard for us to get guidance on what would be a reasonable way to have exogenous borrowing constraints depend on parental income and ability of the child. Hence, we have decided to report a very simple and transparent case in the paper: we assume that children whose both parents have a college degree can borrow twice the amount of the Stafford loan limit. Admittedly, this is ad-hoc in two ways. The first ad-hoc decision is to separate children along the parental education dimension. Our motivation was that parental education strongly correlates with both parental earnings and child's ability. The second ad-hoc decision we faced was: how much more can these children with highly educated parents borrow? We here decided to just double the

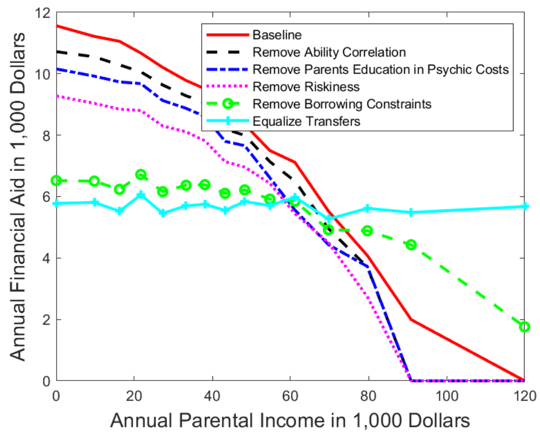


Figure 10: Optimal Financial Aid for Different Model Specifications

Notes: For each model specification (see Figure 9), we illustrate the respective optimal financial aid schedule.

amount in the case that we report. The optimal utilitarian financial aid with parental education dependent borrowing constraints are shown in Figure 12. The shape is slightly different from the baseline optimal schedule, as changes in the borrowing constraints lead to changes share of marginal students.[22] However, the optimal financial aid is still highly progressive.

6.3 Details: Endogenous Ability

We assume that initial ability θ_0 is distributed as:

$$\ln \theta_0 = \beta_0 + \beta_1 \ln I + u$$

where u is normally distributed. We choose β_0, β_1, and the variance of u to match the mean and variance of log childhood ability and covariance of log childhood ability and log parental income from Agostinelli and Wiswall (2016).

We need to calibrate the parameters of the childhood ability production function:

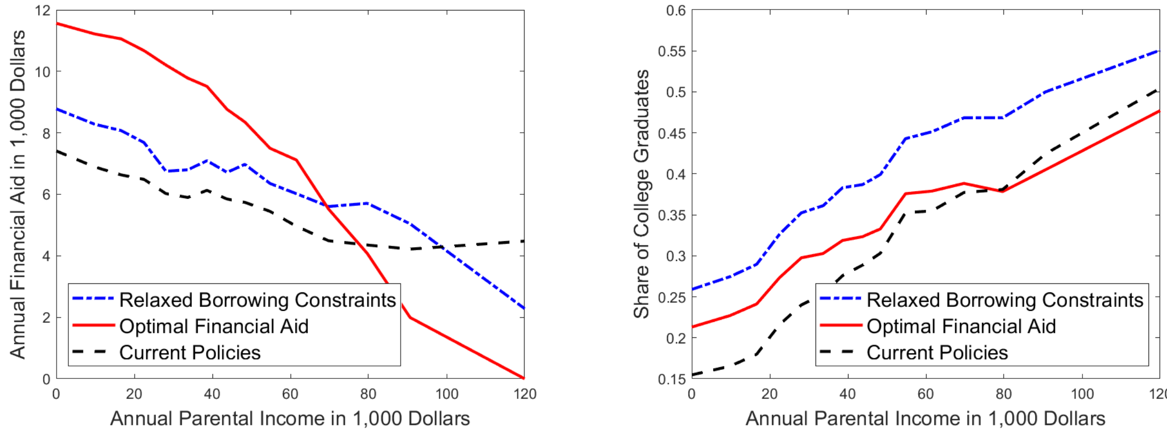
1. A- TFP of parental production function.

2. γ_1 - weight on initial ability

3. γ_2 - weight on parental investment

4. γ_3 - interaction term

[22]As we have shown earlier, relaxing borrowing constraints for all students reduces the progressivity of the optimal aid schedule. That force is still present here, as some low income students have two college educated parents. However, this force is partially muted by the fact that parental education is increasing in parental income. As such, the optimal aid schedule here is more progressive than the case with relaxed borrowing constraints for all individuals, but slightly less progressive than the baseline case with equal borrowing constraints for all students.



(a) Financial Aid

(b) Graduation Rates

Figure 11: Financial Aid and Graduation with Free Borrowing

Notes: The dashed-dotted (blue) line shows the optimal schedule with no borrowing constraints. Optimal financial aid with a Utilitarian welfare function and with borrowing constraints and current financial aid are also shown for comparison in Panel (a). In Panel (b) we display the college graduation share by parental income group for each of the three scenarios.

5. σ^{ι} - variance of ι.

Agostinelli and Wiswall (2016) estimate a translog production function of the following form:

$$\ln \theta_{t+1} = \ln A_t + \gamma_{1t} \ln \theta_t + \gamma_{2t} \ln I_t + \gamma_{3t} \ln \theta_t \cdot \ln I_t + \eta_{\theta,t}$$

for $t = 0, 1, 2, 3$. By combining these four equations, we can derive a single equation for end of childhood ability $\ln \theta_4$ as a function of initial ability $\ln \theta_0$, parental investment in each period $\ln I_t$, the yearly shocks $\eta_{\theta,t}$, and the technology parameters.

Specifically, after some algebra we can write

$$\ln \theta_4 = \ln \theta_0 \left(\gamma_{30} \ln I_0 + \gamma_{10}\right) \left(\gamma_{13} + \gamma_{33} \ln I_1\right) \left(\gamma_{12} + \gamma_{32} \ln I_2\right) \left(\gamma_{11} + \gamma_{31} \ln I_1\right) + f(I, A, \gamma)$$

where $f(I, A, \gamma)$ is a function that depends on investment and the other parameters but not directly on initial ability $\ln \theta_0$.

We can further rearrange this equation to yield

$$\ln \theta_4 = \tilde{\gamma} \ln \theta_0 + g\left(\theta_0, I_0, I_1, I_2, I_3\right) + f(I, A, \gamma)$$

where

$$\tilde{\gamma} = \gamma_{10}\gamma_{11}\gamma_{12}\gamma_{13}$$

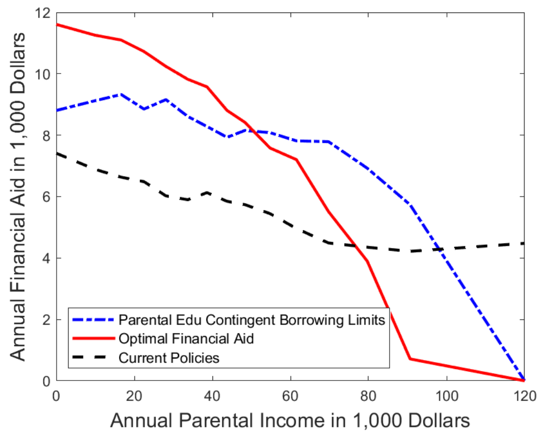


Figure 12: Optimal Financial Aid with Parental Education Dependent Borrowing Constraints

Notes: The dashed-dotted (blue) line shows the optimal schedule with parental income dependent borrowing constraints. Optimal financial aid with in the baseline case and current financial aid are also shown for comparison in Panel (a). In Panel (b) we display the college graduation share by parental income group for each of the three scenarios.

and

$$g\left(\theta_0, I_0, I_1, I_2, I_3\right) = \ln \theta_0 \left(\gamma_{30} \ln I_0 + \gamma_{10}\right)\left(\gamma_{13} + \gamma_{33} \ln I_1\right)\left(\gamma_{12} + \gamma_{32} \ln I_2\right)\left(\gamma_{11} + \gamma_{31} \ln I_1\right) - \tilde{\gamma} \ln \theta_0$$

We set γ_1 equal to the product of the coefficients on lagged ability from Agostinelli and Wiswall (2016) $\gamma_{1,1}\gamma_{1,2}\gamma_{1,3}\gamma_{1,4} \approx 2$. This approximation will be true if the terms on the interaction terms in Agostinelli and Wiswall (2016) are close to zero. Agostinelli and Wiswall (2016) estimate $\gamma_{30} = -0.105(0.066), \gamma_{31} = -0.005(0.019), \gamma_{32} = -0.003(0.013), \gamma_{33} = 0.003(0.010)$. None of the estimates are statistically different from 0 at 95% confidence level and only the first one at a 90% confidence level. Therefore, we think calibrating $\gamma_1 = 2$ seems like a reasonable choice.

Then we have four parameters, A, γ_2, σ^ι, and γ_3. We choose these parameters to match the four following moments:

1. Mean of θ
2. Variance of θ
3. Covariance of θ and parental income I.
4. From Agostinelli and Wiswall (2016): The effect on realized years of schooling of a monetary transfer to parents is roughly ten times larger for parents in the 10th percentile of the income distribution compared to those in the 90th percentile.

Loosely speaking, the covariance of θ and I helps to pin down the importance of parental monetary investments γ_2. The variance of θ helps to pin down the variance to shock of ability production, σ^ι. The differential effect of monetary transfers for rich and poor parents helps to pin down the interaction between parental investment and initial ability, γ_3. Finally, the average ability level helps to discipline the TFP of the production function, γ_1.

Finally, we need to translate these measures of final ability, which are in the units used in Agostinelli and Wiswall (2016) into our measure of ability, which is based on AFQT scores. Let $\hat{\theta}$ represent end of childhood ability as measured in the units used in Agostinelli and Wiswall (2016). We assume that our measures of ability θ is a linear projection of this log skill measure

$$\theta = \alpha_0 + \alpha_1 \ln \hat{\theta}$$

where we choose α_1 and α_0 to match the mean and variance of our AFQT measure. Therefore, when we simulate the model, we first simulate childhood ability in the units used in Agostinelli and Wiswall (2016). Then we translate the measures of ability in Agostinelli and Wiswall (2016) to the ability measures we use in this paper.

Dahl and Lochner (2012) use changes in the EITC to instrument for family income and find that a \$1000 increase in family income leads to an increase in ability scores by 6% of a standard deviation. Increasing yearly family income of parents by \$1,000 in our model leads to an average increase in AFQT scores of 2.2% of a standard deviation across all children, and an increase of 5.1% of a standard deviation for children in the lowest income quintile.

Changes in Childhood Ability Figure 13 shows the change in the relationship between parental income and ability as a result of switching from the current financial aid system to the optimal system with endogenous ability. Ability is measured in percentiles of AFQT scores where percentiles are evaluated at their current levels. We can see that switching to the optimal aid schedule leads to substantial increases in child ability, especially for children in the lower end of the parental income distribution.

6.4 Endogenous Ability with Parental Borrowing Constraints

One issue with the preceding analysis is that we have assumed that parents do not face borrowing constraints. Poor parents may be borrowing constrained while their children are young and therefore may not be able to increase investment in their children in response to changes in financial aid. To explore how borrowing constraints would affect the optimal policy, we assume that P% of parents without a college education cannot increase their investment in their children while the remainder of parents may choose their investment level without this constraint.[23] The optimal policy for a range of values of P is displayed in Figure 14. We can see that the optimal

[23] Caucutt and Lochner (2017) find that 20% of parents with a high school degree and young children are borrowing constrained. Of course, borrowing constraints will also affect the investment decisions of parents who are not at the borrowing limit.

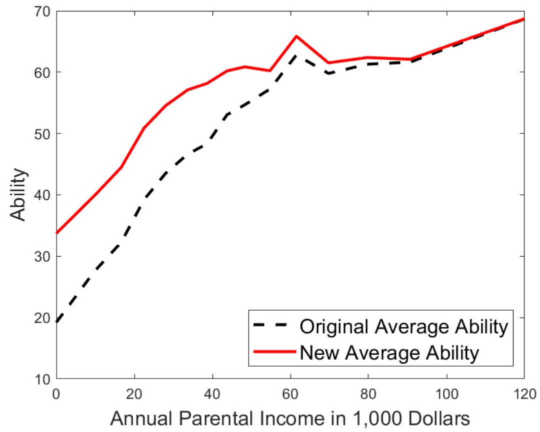


Figure 13: Ability Levels with Endogenous Ability

Notes: This figure shows the relationship between parental income and ability in the optimal system with endogenous ability and under the current financial aid system. Ability is measured in percentiles of the AFQT distribution before financial aid is re-optimized.

progressivity of the system decreases as we increase the percentage of low-education families who are borrowing constrained. However, the optimal schedule remains more progressive than the current schedule in all cases.

6.5 General Equilibrium Effects on Wages

Our analysis abstracted from general equilibrium effects on relative wages. Accounting for these effects would imply that the effects of financial aid on enrollment might be mitigated in the long run: if more individuals go to college, the college wage premium should be expected to decrease because of an increase in the supply of college educated labor (Katz and Murphy, 1992). This in turn would mitigate the initial enrollment increase. To investigate the role of general equilibrium effects on our results, we recalculate the optimal financial aid schedule under the assumption that wages are determined in equilibrium. We assume firms use a CES production function that combines total efficiency units of labor supplied by skilled and unskilled workers, implying that wages are determined by the ratio of skilled to unskilled labor. We assume an elasticity of substitution between skilled and unskilled workers of 2.

We assume identical perfectly competitive firms use CES production functions which combine skilled and unskilled labor. Therefore, wages are determined as a function of the ratio of the total skilled labor to the total unskilled labor.

Let P^U and P^S denote the endogenously determined efficiency wages for unskilled and skilled workers, respectively, where skilled workers are those with a college degree and unskilled workers are high school graduates. We allocate half of college dropouts to each of the skill groups, as is common in the literature (e.g. Card and Lemieux (2001)). Suppose an agent's wages can be written as the product of her efficiency wage and her quantity of efficiency units of labor supplied:

$w_{it} = P^{sk} H_{it}$, where $sk \in \{unskilled, skilled\}$ denotes skill level and H_{it} denotes agent i's level of human capital.[24]

We assume perfectly competitive labor markets. Production at the representative firm is a CES function combining skilled and unskilled labor:

$$Y = A \left(\lambda S^{(\sigma-1)/\sigma} + (1-\lambda) U^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}$$

where A is total factor productivity, λ is the factor intensity of skilled labor, and σ is the elasticity of substitution between skilled and unskilled labor. We assume $\sigma = 2$. S and U represent the total amount of human capital units supplied by skilled and unskilled workers. We assume the economy is in a long run steady-state equilibrium, and that the economy consists of identical overlapping cohorts. Therefore, as cohorts are identical, the total labor supply in the steady-state equilibrium is equal to the total amount of labor supplied over the life-cycle for a given cohort.

Therefore, we can write:

$$S = \sum_i \sum_t H_{it} \ell_{it} \mathbb{I} \left(sk_i = skilled \right)$$

and

$$U = \sum_i \sum_t H_{it} \ell_{it} \mathbb{I} \left(sk_i = unskilled \right)$$

.

Efficiency wages are given by the first order conditions of the firm's profit maximization problem:

$$P^S = A \left(\lambda S^{(\sigma-1)/\sigma} + (1-\lambda) U^{(\sigma-1)/\sigma} \right)^{1/(\sigma-1)} \lambda S^{-1/\sigma}$$

and

$$P^U = A \left(\lambda S^{(\sigma-1)/\sigma} + (1-\lambda) U^{(\sigma-1)/\sigma} \right)^{1/(\sigma-1)} (1-\lambda) U^{-1/\sigma}.$$

These two functions determine wages endogenously as functions of labor supply.

The optimal financial aid schedule and graduation rates with general equilibrium wages are shown in Figures 15(a) and 15(b). We can see that the overall amount of aid has decreased slightly as the fiscal externality of college has been scaled down by general equilibrium wage effects. However, the optimal aid schedule with endogenous wages is just as progressive as in the case with exogenous wages. Thus, while general equilibrium wages dampen the effectiveness of

[24] We normalize units of human capital such that $H_{it} = 1$ is an efficiency unit of labor is defined as the labor supplied by a male worker whose log wages at age 18 are equal to the constant of the wage equation. Therefore, the constants of the wage functions for skilled and unskilled workers are equal to the logs of the efficiency wages for skilled and unskilled workers.

financial aid overall, they do not lead to dramatic changes in the *relative* benefit of financial aid increases for students of different parental income levels. Hence, whereas the overall (average) generosity of the optimal financial aid schedule is slightly lower, the implications for how financial aid should vary with parental income are unchanged.[25]

6.6 Jointly Optimal Financial Aid and Income Taxation

The size of the fiscal externality of college education depends on the tax and transfer system in place. Our structural estimates took the current US tax system as given. An interesting question to ask is how optimal subsidies change when the tax schedule is chosen optimally. To address this, we enrich the optimal policy space such that the planner can also pick a nonlinear tax function $T(y)$ as is standard in the public finance literature (Piketty and Saez, 2013).[26]

Figure 16(a) displays optimal average tax rates in the optimal as well as in the current US system. Average tax rates are higher for most part of the income distribution. As Figure 16(b) shows, this is driven by higher marginal tax rates 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 effect 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 17(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 the 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, strikingly, the progressivity of optimal financial aid policies is preserved. Progressive taxation does not change the desirability of progressive financial aid policies.

6.7 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 18(a) shows that – if optimally targeted also in terms of merit – financial aid policies can

[25]Our results are, hence, consistent with the important earlier paper(s) by Heckman et al. (1998). They find that GE effects dampen the effectiveness of tuition subsidies, and in our case the average level of financial aid is also affected.

[26]We abstract from education dependent taxation; for such cases please see Findeisen and Sachs (2016) and Stantcheva (2017).

be more generous. The progressive nature however is even slightly reinforced. Figure 18(b) shows how optimal financial aid is increasing in AFQT. Interestingly, the slope is almost independent of parental income.

7 Computation of Optimal Policies

We begin by introducing some new notation to ease exposition. We write the current financial aid function as $\mathcal{G}^0(X,I)$. Let $\bar{X}$ denote the vector of median values of X and $\mathcal{G}^0(I)=\mathcal{G}^0\left(\bar{X},I\right)$ be the financial aid schedule faced by agents with the median values of X. In practice, this refers to agents with the median ability. Then we can write the current aid schedule as

$$\mathcal{G}^0(X,I)=\underbrace{\mathcal{G}^0(I)}_{\text{Need Based}}+\underbrace{\hat{\mathcal{G}}^0(X,I)}_{\text{Merit Based}}$$

where we will refer to $\mathcal{G}^0(I)$ as the "need-based" component of the current aid schedule and $\hat{\mathcal{G}}^0(X,I)$ as the "merit-based" component of the current aid schedule, where we normalize $\hat{\mathcal{G}}^0\left(\overline{X},I\right)=0$. For all of the counterfactuals, except when we jointly optimize need and merit-based aid, we will optimize over the need-based component while holding the merit-based component at the current level $\hat{\mathcal{G}}^0(X,I)$. That is, given an alternative need-based component $\mathcal{G}^k(I)$, the financial aid schedule is given by

$$\mathcal{G}^k\left(\bar{X},I\right)=\mathcal{G}^k(I)+\hat{\mathcal{G}}^0(X,I).$$

We will choose the need-based component that solves the government's problem. For the remainder of this section, we will refer to $\mathcal{G}^k(I)$ as the "financial aid schedule" as shorthand for the full financial aid schedule:

$$\mathcal{G}^k(X,I)=\mathcal{G}^k(I)+\hat{\mathcal{G}}^0(X,I).$$

Further, let:

- $\bar{U}^k(I)=\int_{\mathcal{X}}\max\{V^E(X,I),V^H(X,I)\}\tilde{k}(X,I)dX$ denote the Pareto weighted sum of expected lifetime utility of agents with parental income level I given the financial aid schedule $\mathcal{G}^k(I)$.
- $\overline{\mathcal{NT}}^k_{NPV}(I)$ denote the sum of the net-present value of net tax (taxes minus grants) paid by agents with parental income level I given the financial aid schedule $\mathcal{G}^k(I)$.
- For all optimal aid calculations we discretize parental income I. Let $\mathcal{I}$ denote the set of discretized income values. For ease of exposition and with some abuse of notation, we will use $I\in\mathcal{I}$ to denote the discretized values of parental income.

- We will write the government's Lagrange function as

$$\mathcal{L}\left(\mathcal{G}^{k}\left(I\right),\lambda\right)=\sum_{I\in\mathcal{I}}\bar{U}^{k}\left(I\right)-\lambda\left(\sum_{I\in\mathcal{I}}\overline{\mathcal{NT}}^{k}_{NPV}\left(I\right)-\bar{F}\right)$$

7.1 Baseline

1. Calculate the exogenous government revenue requirement. Given the current aid schedule $\mathcal{G}^{0}\left(I\right)$, we calculate the total net tax payments $\sum_{I\in\mathcal{I}}\overline{\mathcal{NT}}^{0}_{NPV}\left(I\right)$. The exogenous government revenue requirement is then given by $\bar{F}=\sum_{I\in\mathcal{I}}\overline{\mathcal{NT}}^{0}_{NPV}\left(I\right)$.

2. Guess a value of the Lagrange multiplier. Denote this guess as $\hat{\lambda}$.

3. Given the current guess of $\hat{\lambda}$, make an initial guess of the need-based aid schedule.

4. Denote the guess of the need-based aid schedule by k. Calculate the sum of utility $\bar{U}^{k}\left(I\right)$ and $\overline{\mathcal{NT}}^{k}_{NPV}\left(I\right)$ for each income level given k.

5. We now perturb the aid schedule to calculate the numerical derivatives of utility and net tax payments with respect to financial aid. Let $\mathcal{G}^{\hat{k}}\left(I\right)=\mathcal{G}^{k}\left(I\right)+\epsilon$. Calculate the sum of utility $\bar{U}^{\hat{k}}\left(I\right)$ and $\overline{\mathcal{NT}}^{\hat{k}}_{NPV}\left(I\right)$ for each income level given $\hat{k}$.

6. Calculate the vector of numerical derivatives of the government's problem with respect to financial aid given multiplier $\hat{\lambda}$, evaluated at $\mathcal{G}^{k}\left(I\right)$. We can write this as:

$$\frac{\Delta\mathcal{L}\left(\mathcal{G}^{k}\left(I\right),\hat{\lambda}\right)}{\epsilon}=\frac{\bar{U}^{\hat{k}}\left(I\right)-\bar{U}^{k}\left(I\right)}{\epsilon}-\hat{\lambda}\frac{\overline{\mathcal{NT}}^{\hat{k}}_{NPV}\left(I\right)-\overline{\mathcal{NT}}^{k}_{NPV}\left(I\right)}{\epsilon}$$

7. Check if each element of $\frac{\Delta\mathcal{L}\left(\mathcal{G}^{k}(I)\hat{\lambda}\right)}{\epsilon}$ equals 0. If not, update the guess of $\mathcal{G}^{k}\left(I\right)$ and go back to step 4. If so, move on to the next step.

8. Given the current guess of $\hat{\lambda}$ and the current guess of the financial aid schedule $\mathcal{G}^{k}\left(I\right)$, check if the government's budget constraint holds with equality ($\bar{F}=\sum_{I\in\mathcal{I}}\bar{\mathcal{NT}}^{k}_{NPV}\left(I\right)$). If the budget does not balance, update the guess of $\hat{\lambda}$ and return to step 3. If the budget balances, then $\mathcal{G}^{k}\left(I\right)$ maximizes the governments problem.

7.2 No taste for redistribution

We follow a similar procedure to the baseline case with one adjustment. With no taste for redistribution, we set the marginal welfare weights to be constant across all individuals. A transfer of ϵ to an inframarginal agent is therefore valued by the social planner as $C\epsilon$, where C is the constant marginal welfare weight. We normalize $C=1$ without loss of generality. Therefore, if the government increases financial aid by ϵ for a given income group I, the increase in social welfare is simply equal to the total discounted years of schooling of agents in income level I. We denote the total discounted years of schooling of agents in income level I under the financial aid schedule indexed by k as $\tilde{E}^k(I)$. Therefore, we again search for the value of the multiplier $\hat{\lambda}$ and the finanicial aid schedule $\mathcal{G}^k(I)$ to maximize the social planner's objective subject to the budget constraint. We perform the same procedure as in the baseline case, except we replace step 6 with the following:

6' Calculate the vector of numerical derivatives of the government's problem with respect to financial aid given multiplier $\hat{\lambda}$, evaluated, at $\mathcal{G}^k(I)$. We can write this as:

$$\frac{\Delta\mathcal{L}\left(\mathcal{G}^k(I)\,\hat{\lambda}\right)}{\epsilon} = \tilde{E}^k(I) - \hat{\lambda}\frac{\bar{\mathcal{N}}\mathcal{T}^{\hat{k}}_{NPV}(I) - \bar{\mathcal{N}}\mathcal{T}^{k}_{NPV}(I)}{\epsilon}$$

Note that, while we set C to be constant across individuals, we still choose the Lagrange multiplier $\hat{\lambda}$ to balance the government's budget constraint. Therefore $\frac{1}{\hat{\lambda}}$ gives the money-metric marginal social welfare weight.

7.3 Revenue Maximizing Government

In this case, the government chooses the financial aid schedule that maximizes $\sum_{I\in\mathcal{I}}\bar{\mathcal{N}}\mathcal{T}^{k}_{NPV}(I)$. We start with an initial guess of the financial aid schedule. Denote this by $\mathcal{G}^k(I)$.

1. Denote the guess of the need-based aid schedule by k. Calculate $\bar{\mathcal{N}}\mathcal{T}^{k}_{NPV}(I)$ for each income level given k.

2. Perturb the aid schedule to calculate the numerical derivative of net tax payments with respect to financial aid. Let $\mathcal{G}^{\hat{k}}(I) = \mathcal{G}^k(I)+\epsilon$. Again calculate the sum of utility $\bar{\mathcal{N}}\mathcal{T}^{\hat{k}}_{NPV}(I)$ for each income level given $\hat{k}$.

3. Calculate the vector of numerical derivatives of the government's problem with respect to financial aid at $\mathcal{G}^k(I)$. This is given by:

$$\frac{\bar{\mathcal{N}}\mathcal{T}^{\hat{k}}_{NPV}(I) - \bar{\mathcal{N}}\mathcal{T}^{k}_{NPV}(I)}{\epsilon}$$

4. Check if each element of $\frac{\bar{\mathcal{N}}\mathcal{T}^{\hat{k}}_{NPV}(I) - \bar{\mathcal{N}}\mathcal{T}^{k}_{NPV}(I)}{\epsilon}$ equals 0. If not, update the guess of $\mathcal{G}^k(I)$ and go back to step 4. If so, then $\mathcal{G}^k(I)$ maximizes the governments problem.

7.4 Alternative Environments

In Section 6.1 we calculate the optimal financial aid policy under the assumption that borrowing constraints are relaxed and in Section 5.2 we calculate the optimal policy under a number of different assumptions about the environment. For each of these alternative environments we perform the same procedure as in the baseline case under the different assumptions about the economic environment.

One point warrants emphasis. Given an alternative environment, we recalculate the exogenous government revenue requirement. That is, letting q index alternative environments, we calculate the exogenous revenue require as $\bar{F}_q = \sum_{I \in \mathcal{I}} \bar{\mathcal{N}\mathcal{T}}^{0}_{NPV,q}(I)$, where $\bar{\mathcal{N}\mathcal{T}}^{0}_{NPV,q}(I)$ is the total net tax payments of agents in income level I, given the environment q and the baseline financial aid schedule. We then calculate the optimal aid schedule given the environment q and the revenue requirement $\bar{F}_q$.

References

Agostinelli, F. and M. Wiswall (2016): "Estimating the Technology of Children's Skill Formation," *NBER Working Paper No. 22442.*

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

Bagnoli, M. and T. Bergstrom (2005): "Log-Concave Probability and Its Applications," *Economic Theory*, 26, 445–469.

Bajari, P., C. S. Chu, D. Nekipelov, and M. Park (2016): "Identification and semiparametric estimation of a finite horizon dynamic discrete choice model with a terminating action," *Quantitative Marketing and Economics*, 14, 271–323.

Card, D. and T. Lemieux (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *The Quarterly Journal of Economics*, 116, 705–746.

Carneiro, P. and J. J. Heckman (2003): "Human Capital Policy," *IZA Discussion Paper No. 821*.

Castleman, B. L. and B. T. Long (2016): "Looking beyond Enrollment: The Causal Effect of Need-Based Grants on College Access, Persistence, and Graduation," *Journal of Labor Economics*, 34, 1023–1073.

Caucutt, E. M. and L. Lochner (2017): "Early and Late Human Capital Investments, Borrowing Constraints, and the Family," Tech. rep.

Dahl, G. B. and L. Lochner (2012): "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit," *American Economic Review*, 102, 1927–1956.

Deming, D. and S. Dynarski (2009): "Into College, Out of Poverty? Policies to Increase the Postsecondary Attainment of the Poor," *NBER Working Paper 15387.*

Diamond, P. A. and E. Saez (2011): "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, 25, 165–190.

Findeisen, S. and D. Sachs (2016): "Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans," *Journal of Public Economics*, 138, 1–21.

Guner, N., R. Kaygusuz, and G. Ventura (2014): "Income Taxation of US Households: Facts and Parametric Estimates," *Review of Economic Dynamics*, 17, 559–581.

Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal Tax Progressivity: An Analytical Framework," *Quarterly Journal of Economics*, 132, 1693–1754.

Heckman, J. J., L. Lochner, and C. Taber (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents U," *Review of Economic Dynamics*, 1, 158.

Hilger, N. G. (2016): "Parental Job Loss and Children's Long-Term Outcomes: Evidence from 7 Million Fathers' Layoffs," *American Economic Journal: Applied Economics*, 8, 247–283.

Katz, L. F. and K. M. Murphy (1992): "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107, 35–78.

Laun, T. and J. Wallenius (2016): "Social Insurance and Retirement: a Cross-Country Perspective," *Review of Economic Dynamics*, 22, 72–92.

Lee, D., S. Y. Lee, and Y. Shin (2017): "The Option Value of Human Capital: Higher Education and Wage Inequality," *Working Paper.*

Matzkin, R. L. (1991): "Semiparametric estimation of monotone and concave utility functions for polychotomous choice models," *Econometrica: Journal of the Econometric Society*, 1315–1327.

——— (1992): "Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models," *Econometrica: Journal of the Econometric Society*, 239–270.

——— (1993): "Nonparametric identification and estimation of polychotomous choice models," *Journal of Econometrics*, 58, 137–168.

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. 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, 205–229.

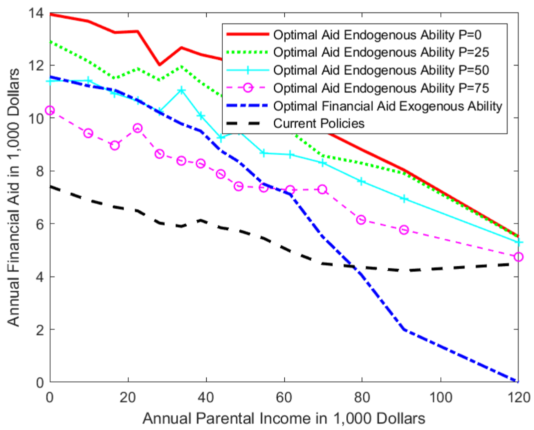
——— (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117, 1039–1073.

Shaked, M. and J. G. Shanthikumar (2007): *Stochastic Orders*, Springer Series in Statistics, New York: Springer.

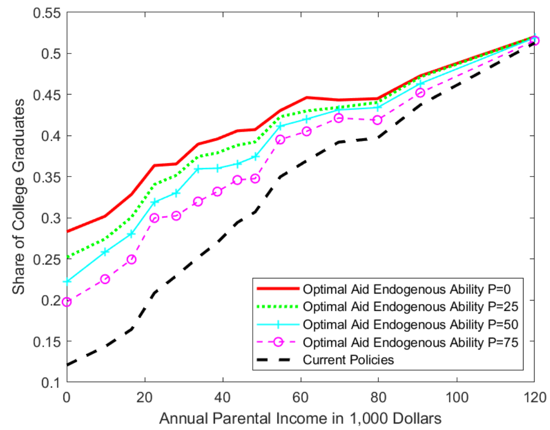
Snyder, T. D. and C. M. Hoffman (2001): *Digest of Education Statistics 2000*, National Center for Education Statistics.

Stantcheva, S. (2017): "Optimal Taxation and Human Capital Policies over the Life Cycle," *Journal of Political Economy*, 125, 1931–1990.

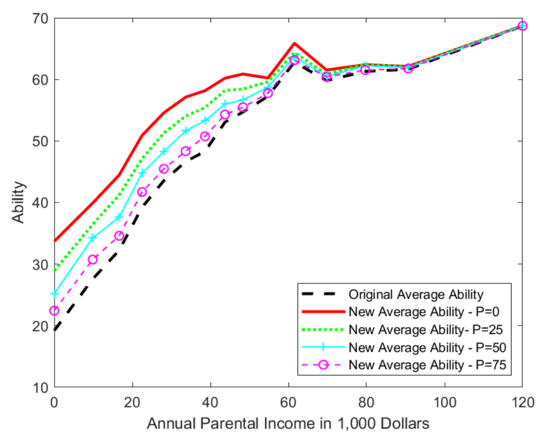
Zimmerman, S. (2014): "The Returns to College Admission for Academically Marginal Students," *Journal of Labor Economics*, 32, 711–754.



(a) Financial Aid



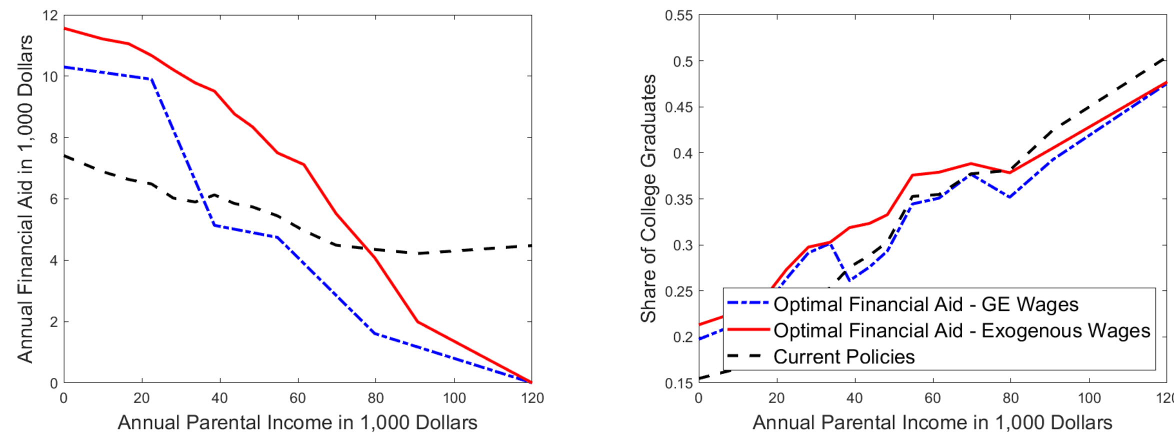
(b) Graduation Rates



(c) Ability Levels

Figure 14: Financial Aid, Graduation and Ability Levels with Endogenous Ability and Parental Borrowing Constraints

Notes: In Panel (a), each line shows the optimal financial aid with endogenous ability when P percent of low-education parents are borrowing constrained and therefore cannot adjust their child's ability in response to changes in financial aid. In Panel (b) we display the college graduation share for each of these scenarios. Panel (c) shows the relationship between parental income and ability in each scenario. Ability is measured in percentiles of the AFQT distribution before financial aid is re-optimized.

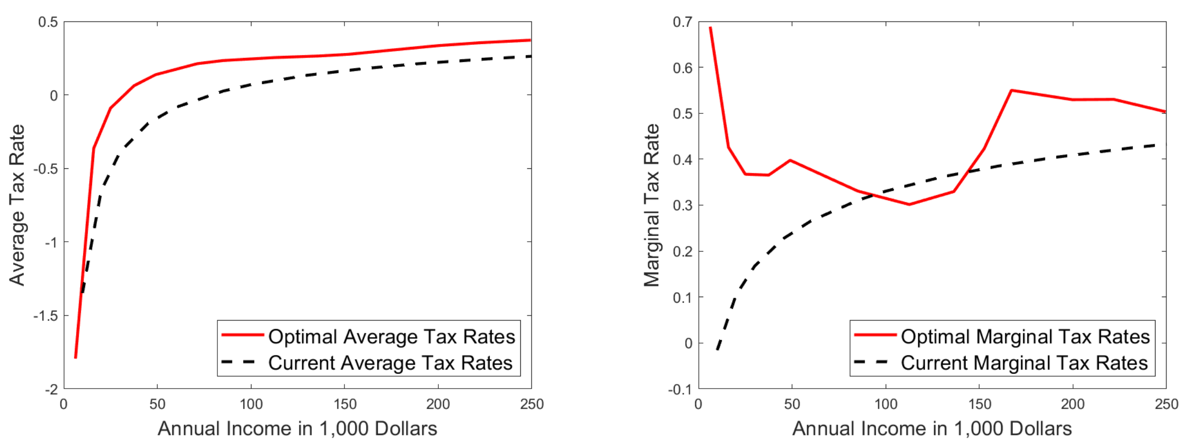


(a) Financial Aid

(b) Graduation Rates

Figure 15: Financial Aid and Graduation with General Equilibrium Wages

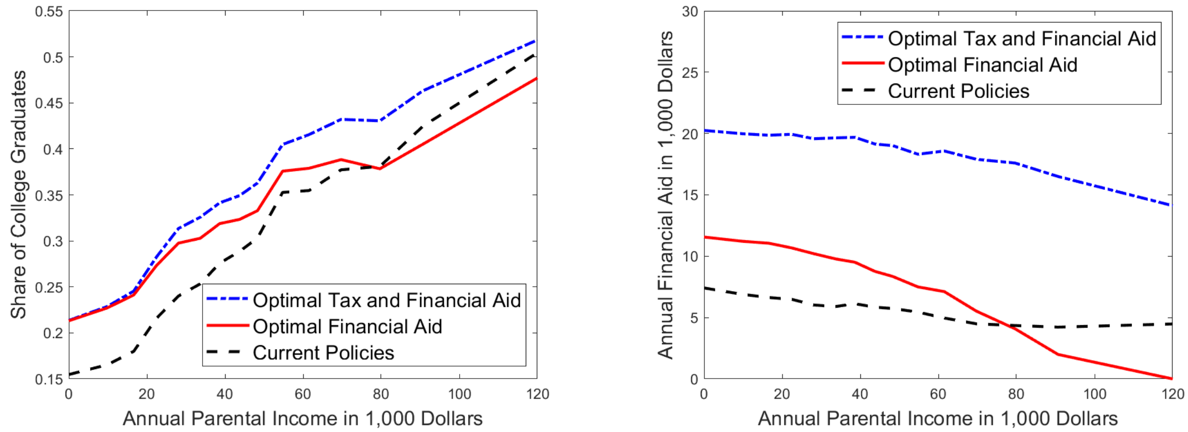
Notes: The dashed-dotted (blue) line shows the optimal schedule when wages are determined in equilibrium. Production is CES between skilled and unskilled workers with an elasticity of substitution of 2. Optimal financial aid with exogenous wage rates and current financial aid are also shown for comparison in Panel (a). In Panel (b) we display the college graduation share by parental income group for each of the three scenarios.



(a) Average Tax Rates

(b) Marginal Tax Rates

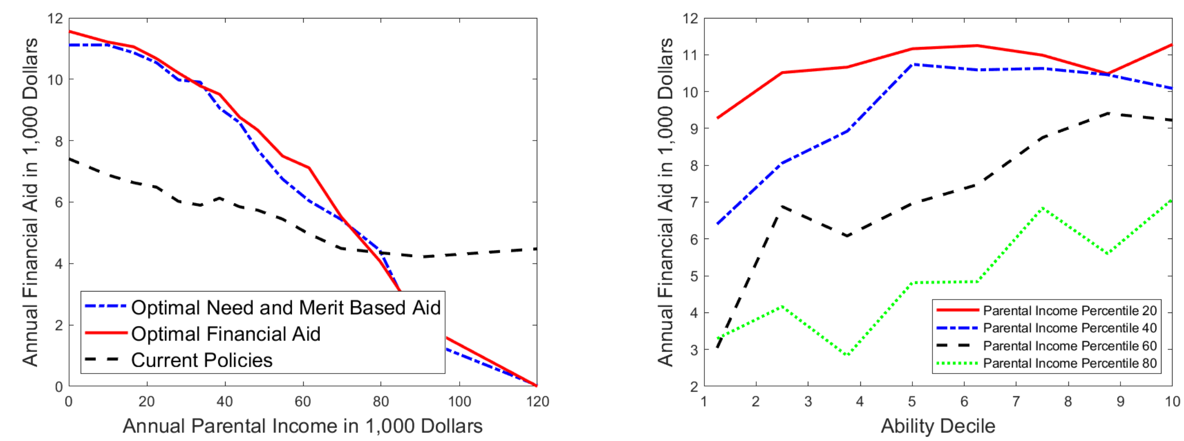
Figure 16: Optimal versus Current: Average and Marginal Tax Rates



(a) Financial Aid (b) Graduation Rates

Figure 17: Financial Aid and Graduation with Optimal Tax Schedule

Notes: The dashed-dotted (blue) line shows the optimal schedule when the tax schedule is also chosen optimally. Optimal financial aid with the current tax schedule and current financial aid are also shown for comparison in Panel (a). In Panel (b) we display the college graduation share by parental income group for each of the three scenarios.



(a) Financial Aid (b) Illustration of Optimal Merit Based Element

Figure 18: Optimal Need and Merit Based Financial Aid

Notes: The dashed-dotted (blue) line shows the optimal financial aid for students with median ability as a function of income when the merit-based component of financial aid is also chosen optimally. Optimal financial aid with exogenous wage rates and current financial aid are also shown for comparison in Panel (a). In Panel (b) the merit based component of the optimal aid schedule.