BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel 2019-7

Inequality in Parental Transfers and Optimal Need-Based Financial Aid

by Youngmin Park

www.bank-banque-canada.ca

Bank of Canada Staff Working Paper 2019-7

February 2019

Last updated September 10, 2025

Inequality in Parental Transfers and Optimal Need-Based Financial Aid

by

Youngmin Park

Canadian Economic Analysis Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
ypark@bankofcanada.ca

ISSN 1701-9397

Acknowledgements

This paper is based on the second chapter of my Ph.D. dissertation at the University of Western Ontario. I am very grateful to my thesis advisors, Lance Lochner, Igor Livshits, and Betsy Caucutt, for their guidance and support. I also thank participants of various seminars and conferences for their valuable feedback. The views expressed in this paper are those of the author and do not necessarily represent those of the Bank of Canada.

Abstract

This paper studies optimal need-based financial aid when parental transfers—unobserved by policymakers—vary across and within families of similar means. Using data on U.S. college students, I document substantial inequality in parental transfers, especially among wealthier families. I then analyze how this affects aid design aimed at reducing inefficiencies from borrowing constraints and the aid itself. Greater inequality in parental transfers among wealthier families weakens the progressivity of the aid, providing more relief to students with low transfers despite their affluent background. Moreover, aid shifts toward two-year colleges, which students with limited parental help are more likely to attend.

Bank topics: Fiscal policy; Potential output; Productivity
JEL codes: D14, D61, D64, D82, I22, J24

Résumé

Cette étude porte sur le fonctionnement optimal de l'aide financière fondée sur les besoins lorsque, sans que les politiques le reflètent, la répartition des transferts parentaux varie à l'intérieur d'une même famille ou entre des familles aux ressources similaires. À l'aide de données sur la population étudiante postsecondaire aux États-Unis, je rends compte de la disparité considérable des transferts parentaux, particulièrement chez les familles mieux nanties. J'analyse ensuite l'effet de cette disparité sur la conception de l'aide financière visant à réduire les inefficiences dues aux contraintes d'emprunt et sur l'aide financière elle-même. Une forte disparité des transferts parentaux parmi les familles mieux nanties affaiblit la progressivité de l'aide financière, puisqu'elle favorise les personnes recevant moins de transferts même si elles sont issues de milieux aisés. De plus, l'aide financière se déplace vers les établissements offrant des programmes de deux ans, qui sont plus susceptibles d'accueillir une clientèle comptant sur une aide financière parentale limitée.

Sujets : Politique budgétaire; Production potentielle; Productivité
Codes JEL : D14, D61, D64, D82, I22, J24

1 Introduction

College education in the U.S. is expensive, and many students rely on financial support from their parents to cover the costs.[1] However, the amount of parental contribution toward college expenses varies widely across students.[2] While some of this variation reflects disparities in families' financial resources (such as parental income and assets), little is known about how parental contributions differ among families with similar economic means.[3]

The current need-based financial aid system in the U.S. adjusts awards based on families' financial capacity, providing more aid to students from households with fewer resources, who tend to receive lower parental transfers.[4] However, if parental support varies substantially even among families with similar financial means, it becomes difficult to precisely target students with limited parental assistance, since parental transfers are not observable or contractible and therefore cannot be directly used in awarding aid. This paper documents considerable inequality in parental transfers among students with comparable family resources and demonstrates that accounting for this variation has important implications for the design of financial aid policy.

Using data from the 2011–2012 National Post-Secondary Student Aid Survey (NPSAS:12), I first document that parental transfers are highly unequally distributed among American undergraduate students considered to have similar family resources for financial aid purposes. These disparities are especially pronounced among wealthier families. Parental transfers also vary across college types conditional on family resources, with contributions generally lower at two-year colleges (which are more affordable) than at four-year colleges.

Motivated by these disparities, I next analytically characterize how differences in the distribution of parental transfers affect the optimal allocation of grant aid (or subsidy) across students with different family resources when they face limited borrowing opportunities.[5] Using a two-period model of binary college attendance choice with heterogeneity in parental transfers, parental income, and preferences for schooling, I show that the optimal aid level for a given parental income group is positively associated with the average marginal utility of consumption among college attendees in that group. Due to diminishing marginal utility, higher parental transfers in higher-income groups lead to lower average marginal utility, and therefore lower

[1]For instance, during the 2023–2024 academic year, the average cost of college for American undergraduates was $28,409, with 49% covered by parents, 27% by grants and scholarships, and 23% by students (Sallie Mae, 2024).

[2]According to the Health and Retirement Survey (HRS), the most common responses regarding the fraction of tuition covered by parents were either 100% or 0% (Haider and McGarry, 2012).

[3]Ellwood and Kane (2000) and Johnson (2013) show a positive correlation between parental income and the amount of financial support provided by parents.

[4]While the role of merit-based aid has grown over time (McPherson and Schapiro, 1998), the majority of non-repayable aid remains need-based: in 2011–2012, only 35% of grant aid received by students at public four-year colleges was non-need-based (College Board, 2016).

[5]Borrowing constraints may stem from frictions such as limited commitment (e.g., Kehoe and Levine, 1993). These frictions are especially relevant for investment in human capital, which serves as poor collateral (Becker, 1975). While the importance of borrowing constraints under current aid policy is debated (see Lochner and Monge-Naranjo (2012) for a review), few would argue they are irrelevant in the absence of financial aid.

optimal aid, under certain distributional assumptions. However, greater inequality in transfers within higher-income groups can offset this effect: when marginal utility is convex, greater dispersion raises the group's average marginal utility, thereby increasing optimal aid, all else equal. These findings—that both the central tendency and dispersion of parental transfers influence the optimal aid schedule—mirror classic results from the literature on decision-making under uncertainty (e.g., Rothschild and Stiglitz, 1970, 1971).

It is important to emphasize that these results are not driven by inequality aversion, which arises from a utilitarian social welfare function with a strictly concave utility function.[6] Instead, I assume that the social planner seeks to minimize inefficiencies in intertemporal consumption allocation and schooling allocation caused by borrowing constraints and the aid policy itself. These inefficiencies are evaluated in monetary terms, based on individuals' willingness to pay to eliminate them.[7] I show that, for a given level of aggregate financial aid spending, minimizing these inefficiencies is equivalent to maximizing a utilitarian social welfare function, where the cardinal scale of the utility function is set in monetary units. While this money-metric utility equals lifetime consumption when borrowing constraints do not bind, it still exhibits concavity and, when individuals are not too averse to shifting consumption over time, also convexity of marginal utility. This is because, when borrowing constraints bind, an additional dollar of consumption during college raises money-metric lifetime utility by more than a dollar, as it improves consumption smoothing over time. Thus, redistributing financial aid toward borrowing-constrained individuals with low parental transfers can be desirable even in the absence of social preferences for equity, because it enhances allocative efficiency.

This framework is extended to allow for multiple college choices—such as the two- versus four-year options considered in the quantitative model—to characterize how optimal aid should vary not only across parental income but also across schooling levels. As in the binary choice case, the optimal aid for a given income and schooling level is positively associated with the average marginal utility of consumption among students in that group. The pattern of average marginal utility across schooling levels, conditional on income, reflects two opposing forces. On the one hand, for a given level of parental transfers, students attending more expensive college options (which also yield higher returns) are more likely to be borrowing constrained and thus have higher marginal utility during school. On the other hand, borrowing constraints affect schooling choices: students who receive lower parental transfers are more likely to select less expensive college options, resulting in a compositional shift toward more constrained students with higher marginal utility at lower schooling levels. This compositional effect is stronger among higher-income families, where parental transfers are more unequal, which in

[6] In contrast to tax/transfer policies, reducing inequality may not be a primary goal of financial aid, which can raise earnings inequality by enabling high-ability students to attend college. For instance, Hanushek, Leung, and Yilmaz (2003) find that education subsidies are less effective at achieving redistribution. Jacobs and Yang (2016) similarly show that optimal education subsidies are lower in the presence of inequality aversion, due to their regressive nature.

[7] Bénabou (2002) develops a related measure of aggregate efficiency that abstracts from equity but does not capture inefficiencies in intertemporal consumption smoothing.

turn weakens the progressivity of aid at lower schooling levels.

These theoretical insights are quantified using a dynastic multi-period lifecycle model with two- and four-year college choice that also accounts for heterogeneity in monetary return to schooling and endogeneous parental transfers. Parental transfers are determined by a family's financial capacity and parents' preferences for giving, captured by the degree of parental altruism, which vary across families.[8] After calibrating parameters related to the monetary costs and returns to schooling, the preference parameters are estimated using data on parental transfers during ages 18–26, education choices, cognitive ability, and other family background variables from the National Longitudinal Survey of Youth 1997 (NLSY97). Importantly, the estimation procedure allows for measurement error in parental transfers by using the value of educational savings accounts intended for the youth's education, measured prior to age 18, as an additional, noisy measure of parental transfers. This additional measurement serves as an instrumental variable, allowing for measurement error to be separately identified from unobserved heterogeneity in parental altruism (Hu, 2008; Hu and Schennach, 2008).

The estimated model is used to compute optimal policies that assign different aid amounts by parental income quartile and college type (two- vs. four-year), approximating the structure of the current U.S. need-based financial aid system. The optimal policy simulations confirm that the theoretical results derived under specific distributional assumptions hold more broadly. The budget-neutral optimal policy yields an aid gap between the bottom and top parental income quartiles at four-year colleges that closely resembles the one observed in the current system. However, this gap would have been larger if parental altruism were homogeneous, eliminating much of the observed dispersion in transfers—especially among higher-income families. The optimal policy also redistributes aid from four-year to two-year colleges across most income quartiles, resulting in a two-year college aid schedule that is flatter and more generous than under current policy. In contrast, under the counterfactual of homogeneous altruism, the optimal policy would have redistributed aid away from two-year colleges.

The optimal policy increases overall college attendance by approximately 2 percentage points and improves efficiency in schooling allocation by about $1,000 per individual (in 2004 U.S. dollars). While it enhances consumption smoothing for students enrolled in two-year colleges, aggregate efficiency in intertemporal consumption allocation remains roughly unchanged, partly because expanded college access introduces consumption distortions for some students who could have fully smoothed consumption had they not attended college.

Related Literature. While many studies quantitatively evaluate higher education policies with parental transfers (e.g., Keane and Wolpin, 2001; Johnson, 2013; Abbott et al., 2019), few examine their optimal design. Conversely, the optimal human capital policy literature

[8] Brown, Scholz, and Seshadri (2012) provide evidence of heterogeneity in parental altruism: in the HRS, 9% of respondents would not give 5% of their income to a child earning only one-third of their own income, while 62% would even if the child earned 75% of their income.

often abstracts from intergenerational linkages (e.g., Bohacek and Kapicka, 2008; Findeisen and Sachs, 2016; Stantcheva, 2017) or considers only policies that do not vary with family background (Krueger and Ludwig, 2016; Jacobs and Yang, 2016).

One exception is Colas, Findeisen, and Sachs (2021), who study how optimal financial aid varies with parental income using an estimated model with endogenous parental transfer decisions. Their work is most closely related to this paper but differs in three important respects. First, they do not examine how the distribution of parental transfers, conditional on income, affects the optimal aid schedule. Second, they focus exclusively on four-year colleges and do not consider how optimal aid varies across college types. Third, in their framework, the primary rationale for financial aid is to offset the distortions from income taxation on education investment (Bovenberg and Jacobs, 2005). This paper complements their work by motivating need-based financial aid as a response to inefficiencies from borrowing constraints, and by examining how aid design is affected when those in need cannot be easily identified due to heterogeneity in parental preferences. As discussed earlier, conditioning aid on both college type and income is a key policy lever for effectively targeting students who are borrowing constrained because of low parental support.

I also extend the analysis to incorporate inefficiencies in schooling allocation arising from social returns that are not internalized by individuals—specifically, the impact of education on future tax revenues. Consistent with Colas, Findeisen, and Sachs (2021), the optimal aid schedule for four-year colleges becomes more progressive than the current policy, expanding aid only for lower-income groups whose stronger enrollment responses generate greater fiscal returns. Nonetheless, the central finding of this paper—the role of the distribution of parental transfers in shaping aid amounts across both parental income and college types—remains. Optimal aid schedules for four-year colleges would be even more progressive under homogeneous parental preferences. Moreover, aid for two-year colleges expands across all income groups relative to the current policy, whereas under homogeneous altruism, the opposite pattern would emerge.

Providing larger subsidies for higher levels of education investment—a feature of current U.S. financial aid policy, which offers more aid to students at four-year than at two-year colleges—is often justified as a way to alleviate borrowing constraints (De Fraja, 2002; Jacobs and Yang, 2016). This paper shows that borrowing constraints can support such a policy when they bind more tightly for students pursuing more expensive options. However, when differences in investment primarily stem from unobserved variation in available resources rather than returns, borrowing constraints may instead justify greater aid at lower levels of schooling. Accounting for heterogeneity in both educational returns and parental support, the quantitative analysis shows that optimal aid for two-year colleges can exceed that for four-year colleges, particularly among high-income families where parental transfers are more unequal.

The rest of the paper is organized as follows. Section 2 presents U.S. evidence on the relationship between parental transfers and family resources. Section 3 uses a two-period model to analytically characterize the optimal financial aid. Section 4 extends the model to a multi-period life-cycle with endogenous parental transfers and describes its parameterization. Section 5 quantitatively characterizes the optimal financial aid. Section 6 further extends the analysis to account for the impact of education on tax revenues. Section 7 concludes.

2 Parental Transfers and Family Resources

In this section, I examine inequality in parental transfers across families with similar resources using data from the NPSAS:12. The NPSAS:12 is a nationally representative survey of students enrolled in U.S. post-secondary institutions during the 2011–2012 academic year. A key feature of this dataset is that it includes both information on family resources—measured by the Expected Family Contribution (EFC)—and on how much parents actually contributed toward college expenses.

The EFC is intended to reflect the amount a student's family is reasonably expected to contribute toward college costs for the academic year.[9] It is calculated based on the information reported in the federal financial aid application, including income, assets, and family size. For dependent students—which includes most traditional college-age students—the EFC is determined primarily by parental income from the previous year, while student income and family net worth play a more limited role.[10][11]

The EFC plays a central role in the allocation of financial aid in the U.S., where most aid is need-based. For federal aid, a student's financial "need" is defined as the difference between the cost of attendance and the EFC.[12] The cost of attendance is determined by each institution and includes tuition, fees, and estimated living expenses. In general, the amount of aid awarded increases with student need (Dick and Edlin, 1997).

In the NPSAS:12 data, the EFC and other financial aid variables are drawn from administrative records, whereas parental transfer amounts are reported by students in the survey. To ensure access to administrative data, I restrict the sample to dependent students who submitted a federal financial aid application. Because cost of attendance can vary widely across institutions and student circumstances, I focus on a more homogeneous group by limiting the sample to U.S. citizens enrolled full time for more than nine months at a single two- or four-year public institution during the 2011–2012 academic year, paying the standard "in-jurisdiction" tuition rate and not living with their parents while enrolled.

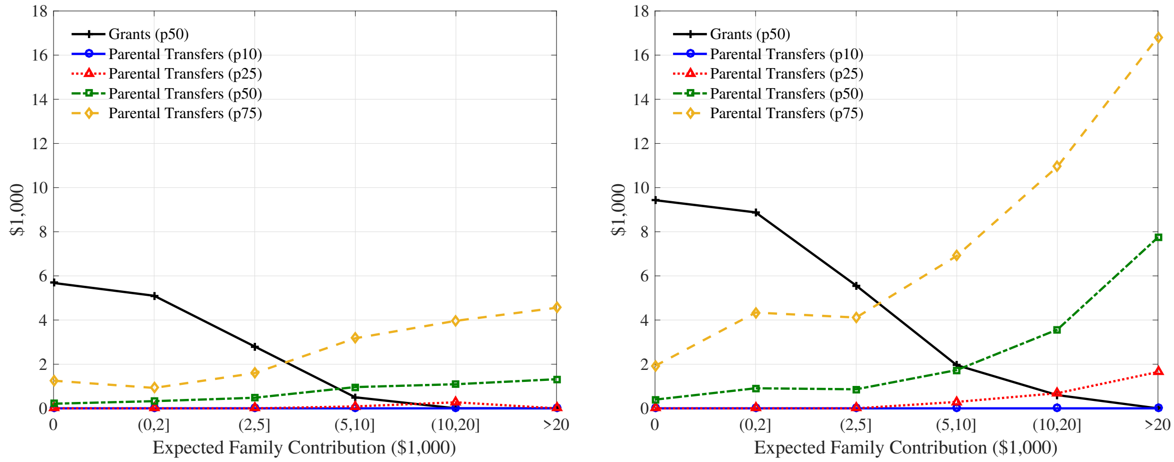
Figure 1 displays percentiles of parental contributions during the 2011–2012 academic year

[9] See Part F of Title IV of the Higher Education Act for further details on the needs analysis formulas.

[10] Students are considered independent if they meet any of the following criteria: they are age 24 or older; married; have dependents; are veterans or active-duty members of the U.S. armed forces; or are orphans or wards of the court.

[11] The EFC calculation excludes assets such as home equity in a primary residence and retirement savings. Liabilities like mortgages on the primary residence and consumer debt are also generally excluded.

[12] State and institutional aid programs generally follow a similar need-based formula.



(a) 2-Year (b) 4-Year

Figure 1: Grants and Parental Transfers by Expected Family Contribution

by six EFC groups, separately for students attending two- and four-year institutions.[13] The figure also includes median grant amounts—the primary subsidy component of financial aid—for each EFC group. As expected under a need-based system, students with higher EFCs receive smaller grants. Correspondingly, parental contributions generally rise with EFC, compensating for the reduction in grant aid.

However, parental contributions vary widely even within each EFC group. For example, over 10% of students with an EFC above $20,000 receive no parental support, while more than 25% of students with a zero EFC receive over $1,000. The dispersion of parental transfers also increases with EFC. At four-year institutions, the interquartile range of parental transfers grows more than sevenfold from the lowest to the highest EFC group. A similar, though more muted, increase is observed at two-year institutions. This widening dispersion is partly driven by a stagnant lower tail: the 25th percentile increases by less than $2,000 across EFC groups at four-year institutions and remains flat at zero for two-year institutions.

Comparing two-year and four-year institutions, students at two-year colleges receive smaller grants than their four-year counterparts within the same EFC group, reflecting the lower cost of attendance. At any given EFC level, parental transfers are also lower for students attending two-year colleges. This may reflect that parents contribute less when their children attend less expensive institutions, but it could also indicate that students anticipating limited parental support are more likely to choose lower-cost options.

Taken together, this evidence suggests that even parents with sufficient means often contribute less than what is expected by financial aid authorities. Given limited borrowing opportunities, their children—who receive not only low parental support but also reduced financial aid—may face low consumption during college or opt for less schooling than they otherwise

[13]See Section A of the Appendix for additional details on the data and summary statistics.

would.[14] Therefore, the financial aid system design must account not only for differences in family resources, but also for variation in parental contributions among families with similar means—a disparity that is particularly pronounced among wealthier families.

3 Theory of Optimal Need-Based Financial Aid

In this section, I develop a simple two-period model to analytically characterize the optimal financial aid schedule when parental transfers are imperfectly correlated with family resources. I begin with a binary schooling choice—college attendance versus non-attendance—and examine how the distribution of parental transfers shapes the optimal progressivity of financial aid, defined as the extent to which aid amounts vary with family resources. I then extend the model to allow for multiple college options with differing costs and returns, highlighting how optimal aid allocation also depends on college choices conditional on family resources.

3.1 Individual's Problem

Intertemporal Consumption Allocation. Consider individuals who may invest in schooling in the first period and work in the second period. Their preferences are represented by the following utility function:

$$U(c_1, c_2) := (1+\beta)\left(\frac{1}{1+\beta}c_1^{\frac{\gamma-1}{\gamma}} + \frac{\beta}{1+\beta}c_2^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}}, \tag{1}$$

where $c_t \geq 0$ is consumption in periods $t \in \{1,2\}$, β is a discount factor, and $\gamma \geq 0$ is the elasticity of intertemporal substitution (EIS).

The utility function (1) is a monotonic transformation of the standard additively separable utility function $u(c_1) + \beta u(c_2)$, where $u(c) = (c^{1-1/\gamma} - 1)/(1 - 1/\gamma)$. As shown in Section B.1 of the Appendix, the transformation converts utils to dollars. This "money-metric" utility function (McKenzie, 1957) is useful for welfare analysis because it eliminates inequality aversion typically present in the utilitarian social welfare function.

Each individual is endowed with a parental transfer $b \geq 0$, which can be thought of as initial wealth. Individuals face two schooling options: $j = 0$ and $j = 1$ represent the option of not attending and attending college, respectively. Let y_j be lifetime earnings associate with schooling option j.

Individuals who do not attend college enter the workforce immediately in the first period and can borrow or save freely at the gross interest rate $R = 1/\beta$. They maximize utility given by equation (1), subject to the lifetime budget constraint $c_1 + R^{-1}c_2 \leq y_0 + b$. Let $V_0(b)$ denote

[14] In 2011–2012, first-year dependent students could borrow only up to \$5,000 in federal loans, well below the average cost of attendance exceeding \$10,000. Most private loans require evidence of creditworthiness, which many undergraduates lack, and carry higher interest rates (Consumer Financial Protection Bureau, 2012).

their indirect utility function. It is easy to see that $V_0(b) = y_0 + b$; that is, lifetime utility in monetary terms equals total lifetime consumption. Consequently, the marginal value of wealth is constant: $V_0'(b) = 1$.

On the other hand, those who attend college pay the monetary cost of college (i.e., tuition) $k \geq 0$, receive financial aid $g \geq 0$, and cannot borrow more than the amount $\bar{l} > k$ in the first period. Let $V_1(b+g)$ be their indirect utility function, that is, maximized utility subject to the lifetime budget constraint $c_1 + R^{-1}c_2 \leq y_1 + b + g - k$ as well as the borrowing constraint

$$c_1 \leq b + g - k + \bar{l}. \tag{2}$$

Note that the assumption $\bar{l} > k$ ensures that consumption during college is positive, making college attendance feasible regardless of the level of parental transfers or financial aid.

When individuals have enough resources ($b + g \geq Ry_1 + k - (1+R)\bar{l}$), the borrowing constraint (2) does not bind and full consumption smoothing is achieved ($u'(c_1) = \beta u'(c_2)$). Otherwise, (2) binds and intertemporal consumption allocation is distorted ($u'(c_1) > \beta u'(c_2)$). In this case, $V_1(b+g) < y_1 + b + g - k$ and $V_1'(b+g) > 1$. That is, the money-metric lifetime utlity is less than lifetime consumption due to the intertemporal consumption distortion, and an additional dollar of initial wealth increases utility by more than a dollar by reducing the consumption distortion. This results is state in the following lemma.[15]

Lemma 1 (Marginal Utility)**.** $V_1'(b+g) \geq 1$*, with strict inequality if and only if* (2) *binds.*

College Attendance Decision. Individuals with an identical parental transfer may make different schooling choices because of differences in non-pecuniary, or "psychic," costs of schooling.[16] Let ε be the utility gain from not attending college, distributed with a cumulative distribution function $F(\cdot)$ and a density function $f(\cdot)$. Then, an individual attends college if and only if $\varepsilon \leq V_1(b+g) - V_0(b)$ and the college attendance rate conditional on (b, g) is

$$p(b,g) := \Pr\left(\varepsilon \leq V_1(b+g) - V_0(b)\right) = F\big(V_1(b+g) - V_0(b)\big).$$

The following lemma characterizes the college attendance decision.

Lemma 2 (College Choice)**.** *(i) $p(b,g)$ is increasing*[17] *in b when* (2) *binds, and is constant in b otherwise; (ii) $p(b,g)$ is increasing in g.*

College attendance increases with parental transfers only when the borrowing constraint (2) binds. In this case, additional wealth raises the value of attending college more than the value of not attending by improving consumption smoothing during college (i.e., $V_1'(b+g) > V_0'(b) = 1$).

[15] See Section B of the Appendix for all proofs and other analytical details.
[16] "For a review of the evidence on non-pecuniary returns to schooling, see Heckman, Lochner, and Todd (2006).
[17] Throughout the paper, "increasing" and "decreasing" refer to weak monotonicity unless otherwise specified.

When the constraint does not bind, parental transfers have no effect on schooling choices because individuals select the option that maximizes lifetime consumption, accounting for psychic costs.[18] This result is well established in the literature (e.g., Becker, 1975; Lochner and Monge-Naranjo, 2011).

In contrast, college attendance increases with higher financial aid regardless of whether the borrowing constraint binds, since aid raises the value of attending without affecting the value of not attending. For borrowing-constrained students, additional aid helps smooth consumption over time, bringing their enrollment decision closer to what it would be in the absence of the constraint. However, because aid is conditional on enrollment, it also introduces a distortion—inducing some students to attend college who otherwise would not.

The adverse effects of borrowing constraints on intertemporal consumption smoothing and educational attainment are central to the optimal design of financial aid in this paper. As noted earlier, while financial aid can alleviate these constraints, it also introduces distortions by influencing schooling choices. Next, I outline how these inefficiencies can be quantified in monetary terms.

3.2 Efficiency Losses

In this subsection, I develop monetary measures of inefficiencies, based on individuals' maximum willingness to pay to eliminate them.

Intertemporal Consumption Distortion. Recall that the money-metric utility function (1) satisfies that $U(c_1, c_2)$ is equal to lifetime consumption $c_1 + R^{-1}c_2$ if $c_1 = c_2$, and strictly less than $c_1 + R^{-1}c_2$ if $c_1 \neq c_2$, reflecting the efficiency loss from the intertemporal consumption distortion. Because individuals are indifferent between a potentially distorted consumption profile (c_1, c_2) and an undistorted consumption profile (c_1^*, c_2^*), where $c_1^* = c_2^* = U(c_1, c_2)/(1+\beta)$, they are willing to pay as much as $c_1 + R^{-1}c_2 - (c_1^* + R^{-1}c_2^*) = c_1 + R^{-1}c_2 - U(c_1, c_2)$ to eliminate the intertemporal consumption distortion.

Formally, the distortion in intertemporal consumption allocation for an average individual receiving (b, g) is defined as:

$$L_c(b, g) \coloneqq p(b, g)\big[(y_1 + b + g - k) - V_1(b + g)\big].$$

When the individual attends college, she consumes $y_1 + b + g - k$ and attains utility $V_1(b + g)$ over the lifetime. However, she could have attained the same level of utility if she consumed a constant amount $V_1(b+g)/(1+\beta)$ in each period. Therefore, $(y_1 + b + g - k) - V_1(b+g)$ is her

[18] This absence of direct wealth effects is a desirable feature, as wealth effects would imply a counterfactual negative relationship between parental resources and educational attainment for empirically relevant values of positive psychic costs (Caucutt, Lochner, and Park, 2017).

maximum willingness to pay to eliminate the intertemporal consumption distortion associated with college attendance. Since the consumption allocation of those who do not attend college is not distorted, taking the average across individuals over the psychic cost ε yields $L_c(b, g)$.

Schooling Distortion. Next, the distortion in schooling allocation conditional on (b, g) is

$$L_s(b, g) := \int \max\{y_0 + \varepsilon, y_1 - k\} - \left\{\mathbb{I}_{\varepsilon > V_1(b+g) - V_0(b)}(y_0 + \varepsilon) + \mathbb{I}_{\varepsilon \leq V_1(b+g) - V_0(b)}(y_1 - k)\right\} dF(\varepsilon),$$

where $\mathbb{I}_x = 1$ if the statement x is true and $\mathbb{I}_x = 0$ otherwise. The first term in the integral is the lifetime income, net of psychic and monetary costs of schooling, delivered by the "first-best" schooling choice, that is, in the absence of borrowing constraints and financial aid. The second term in the integral is the same object, but under the current schooling choice. Individuals are willing to pay the difference between the two if they could eliminate the inefficiencies leading to the current schooling choice that deviates from the first-best level.

Total Distortion. Finally, the total distortion is $L(b, g) := L_c(b, g) + L_s(b, g)$. Aggregating $L(b, g)$ across individuals gives the social planner's objective function to be minimized. The next lemma shows that this monetary measure of inefficiency is closely related to the monetary measure of welfare, $V(b, g) := \int \max\left\{V_0(b) + \varepsilon, V_1(b + g)\right\} dF(\varepsilon)$.

Lemma 3 (Welfare Decomposition)**.** *The average indirect utility can be written as follows:*

$$V(b, g) = \underbrace{\int \max\{y_0 + \varepsilon, y_1 - k\} dF(\varepsilon) + b}_{\textit{first-best}} + \underbrace{p(b, g)g}_{\substack{\textit{gain if aid is given as}\\ \textit{a lump-sum transfer}}} - \underbrace{L(b, g)}_{\textit{efficiency loss}}. \tag{3}$$

The monetary measure of welfare, $V(b, g)$, consists of the maximum lifetime (money-metric) utility that can be attained in the absence of borrowing constraints and financial aid, $\int \max\{y_0 + \varepsilon, y_1 - k\} dF(\varepsilon) + b$, plus the potential increase in utility if the current amount of aid is given instead as a lump-sum transfer, $p(b, g)g$, net of the efficiency losses in intertemporal consumption and schooling allocations, $L(b, g)$.

Since $V(b, g)$ is an indirect utility function, aggregating it across individuals gives a utilitarian social welfare function without inequality aversion. The decomposition (3) implies that maximizing this utilitarian social welfare function is equivalent to minimizing aggregate efficiency losses when the aggregate spending on financial aid is held constant. However, as discussed in the next subsection, the two objectives may lead to different allocations.

3.3 Optimal Policy

Now, I characterize how the optimal financial aid schedule—that minimizes efficiency losses—varies across family resource levels when parental transfers are imperfectly correlated with those resources. Consider individuals who differ in parental transfers, psychic costs of schooling, and parental income, where the latter influences only the distribution of transfers. Specifically, suppose there are finitely many parental income groups indexed by $h \in \mathcal{H}$. Let n_h denote the fraction of individuals in group h, and $m_h(\cdot)$ the probability density function of parental transfers within that group.

Define the college attendance rate for group h, $P_h(g)$, and the density of parental transfers conditional on college attendance within group h, $m_{h|1}(b, g)$, as:

$$P_h(g) := \int p(b, g) m_h(b) db, \quad m_{h|1}(b, g) := \frac{p(b, g) m_h(b)}{P_h(g)}.$$

The social planner, endowed with a total budget $G > 0$, observes each individual's schooling decision and group membership but not their specific parental transfers or psychic costs. The planner sets a fiancial aid schedule to minimize aggregate efficiency losses. Let $g_h \geq 0$ denote the aid amount assigned to group h. The planner's problem is formulated as follows:

Definition 1 (Optimal Policy)**.** *The optimal policy* $(\hat{g}_h)_{h \in \mathcal{H}}$ *solves*

$$\begin{aligned} \min_{(g_h)_{h \in \mathcal{H}}} & \sum_{h \in \mathcal{H}} n_h \int L(b, g_h) m_h(b) db \\ \textit{subject to} & \sum_{h \in \mathcal{H}} n_h P_h(g_h) g_h \leq G, \end{aligned} \tag{4}$$

and $g_h \geq 0$ *for all* $h \in \mathcal{H}$*. Let* $\lambda \geq 0$ *be the Lagrange multiplier on the constraint* (4).

Optimality Condition. The first-order condition with respect to g_h is

$$\underbrace{\int V_1'(b + \hat{g}_h) m_{h|1}(b, \hat{g}_h) db}_{\text{average marginal utility of attendees}} \leq (1 + \lambda) \underbrace{\left[\overbrace{1}^{\text{mechanical}} + \overbrace{\frac{P_h'(\hat{g}_h)}{P_h(\hat{g}_h)} \hat{g}_h}^{\text{behavioral response}} \right]}_{\text{impact on public spending}}, \tag{5}$$

with strict inequality if and only if $g_h \geq 0$ binds. The left-hand side of (5) represents the marginal benefit of financial aid, while the right-hand side captures its marginal cost.

As shown by Colas, Findeisen, and Sachs (2021), an increase in financial aid acts like a lump-sum transfer for inframarginal students—those who would attend college regardless—while it has distortionary effects on marginal students—those who enroll only because of the additional aid.[19] For inframarginal students, the aid increase simply raises their utility, as captured by the average marginal utility, and imposes a direct cost on the planner, referred to as the mechanical effect. When some inframarginal students are borrowing constrained, the utility gain from aid exceeds one, reflecting improved consumption smoothing.

[19] Saez (2001) discusses a similar mechanism in the context of optimal income taxation.

In contrast, marginal students are indifferent between attending and not attending; their enrollment does not increase aggregate utility but still incurs a cost through the rise in attendance. This efficiency loss stems from the behavioral response to the aid expansion across the entire group, since aid cannot be conditioned on parental transfers or psychic costs—factors unobserved by the social planner. Importantly, this inefficiency persists even when college attendance falls short of the first-best level due to borrowing constraints.

To illustrate how minimizing inefficiency differs from standard welfare maximization, consider a group in which no individual faces a binding borrowing constraint. Under an inefficiency-minimizing policy, this group receives no aid, since it would only reduce efficiency. In contrast, welfare maximization may still allocate aid to this group if it raises overall utility. If the inefficiency-minimizing solution does allocate aid to such a group, it implies that their enrollment decisions are unresponsive to aid changes and that the planner's budget constraint (4) is slack.

Proposition 1 (No Aid for the Unconstrained)**.** *Suppose that the borrowing constraint* (2) *is slack for all individuals in group h under the optimal policy. (i) If* $P'_h(\hat{g}_h) > 0$*, then* $\hat{g}_h = 0$*; (ii) If* $\hat{g}_h > 0$*, then* $P'_h(\hat{g}_h) = \lambda = 0$*.*

Moreover, unlike under welfare maximization, the planner's budget constraint can be slack in the inefficiency-minimizing solution. This occurs when no individuals in the economy face binding borrowing constraints, so additional aid does not reduce inefficiencies.

Corollary 1 (Slack Social Budget Constraint)**.** *If the borrowing constraint* (2) *is slack for all individuals, then* $\lambda = 0$*.*

Rearranging the optimality condition (5) gives the optimal policy formula for $\hat{g}_h > 0$:

$$\hat{g}_h = \frac{P_h(\hat{g}_h)}{P'_h(\hat{g}_h)} \left\{ \frac{\int V'_1(b + \hat{g}_h) m_{h|1}(b, \hat{g}_h) db}{1 + \lambda} - 1 \right\}. \tag{6}$$

This expression shows that variation in optimal aid across groups is driven by differences in (i) average marginal utility, which captures the marginal benefit of financial aid, and (ii) enrollment responses, which determine the marginal cost. I examine each component in turn.

Variation of Marginal Utility Across Groups. Parental income groups differ only in the distribution of parental transfers. I focus on two key aspects of these distributions—their central tendency and dispersion—and examine how they influence the average marginal utility. As is well established in the literature on decision-making under uncertainty (e.g., Rothschild and Stiglitz, 1970, 1971), the first and second derivatives of marginal utility determine how its average responds to shifts in the location and spread of the distribution.

Lemma 4 (Derivatives of Marginal Utility)**.** *(i)* $V''_1(b+g) \leq 0$*; (ii)* $V'''_1(b+g) \geq 0$ *if* $\gamma \geq 1/2$*.*

Marginal utility decreases with initial wealth: it is strictly decreasing in the constrained region, where the borrowing constraint binds, and constant—equal to one—in the unconstrained region. As a result, the slope of marginal utility transitions from negative at low wealth to zero at high wealth. This transition is monotonic—and thus implies global convexity of marginal utility—when the EIS is not too low.

Now, I consider two groups h and h', and compare their average marginal utility when their parental transfer distributions differ in either location or spread of the distribution.

Lemma 5 (Greater Parental Transfers)**.** *Suppose that $m_{h'}(b)/m_h(b)$ is increasing in b for all b. Then $\int V_1'(b+g)m_{h'|1}(b,g)db \leq \int V_1'(b+g)m_{h|1}(b,g)db$ for all g.*

If individuals in group h' tend to receive higher parental transfers than those in group h in the sense of monotone likelihood ratio dominance (Milgrom, 1981), then their distribution of transfers conditional on college attendance also shifts upward, exhibiting first-order stochastic dominance. Because students with higher parental transfers are less likely to face binding borrowing constraints—and therefore have lower marginal utility—the average marginal utility among college attendees is lower in group h'.

Lemma 6 (Unequal Parental Transfers)**.** *Suppose that $m_h(b)/m_{h'}(b)$ is unimodal in b and $\int b\left[m_{h|1}(b,g) - m_{h'|1}(b,g)\right]db = 0$. If $\gamma \geq 1/2$, then $\int V_1'(b+g)m_{h'|1}(b,g)db \geq \int V_1'(b+g)m_{h|1}(b,g)db$.*

If group h has a more equal transfer distribution than group h', in the sense of unimodal likelihood ratio dominance (Ramos, Ollero, and Sordo, 2000; Hopkins and Kornienko, 2004), then the distribution of transfers among college attendees in group h is also less dispersed.[20] Moreover, if the two groups of attendees have the same mean parental transfer, the distribution for group h' is a mean-preserving spread of that for group h. Given the convexity of marginal utility, this implies that the average marginal utility among attendees is higher in group h'.

Overall, these results suggest that the marginal benefit of aid is generally lower for higher parental income groups, which tend to have higher average transfers. However, this effect may be attenuated by the greater within-group dispersion in transfers observed among these groups.

Behavioral Response Under a Distributional Assumption. When $p(b,g) \in (0,1)$ for all b, the semi-elasticity of college attendance with respect to financial aid is given by:

$$\frac{P_h'(g)}{P_h(g)} = \int V_1'(b+g)\frac{f\big(V_1(b+g) - V_0(b)\big)}{F\big(V_1(b+g) - V_0(b)\big)}m_{h|1}(b,g)db, \tag{7}$$

[20]The unimodal likelihood ratio dominance is a refinement of second-order stochastic dominance, analogous to how monotone likelihood ratio dominance refines first-order dominance. See Costinot and Vogel (2010) for another application of this concept.

which expresses the behavioral response as the average marginal utility weighted by the reverse hazard rate of the psychic cost distribution, $f(\varepsilon)/F(\varepsilon)$.

Although higher-income groups are likely to exhibit smaller behavioral responses due to their higher attendance rates, comparing these responses across groups requires an assumption about the distribution of psychic costs. Without loss of generality, I model the psychic cost ε as a location-scale transformation of a baseline random variable z, where $\mathrm{E}[z]$ and $\mathrm{SD}(z)$ denote its mean and standard deviation, and impose the following distributional assumption on z:

Assumption 1 (Psychic Cost)**.** *$\varepsilon = \mu + \sigma(z - E[z])/SD(z)$, where $z \leq 0$ follows a distribution with cumulative distribution function e^z, and $\sigma > 0$.*

Under this assumption, the reverse hazard rate simplifies to a constant: $f(\varepsilon)/F(\varepsilon) = 1/\sigma$, over the support $\varepsilon \leq \mu + \sigma$, where μ and σ are the mean and standard deviation of ε. This implies that the semi-elasticity of attendance with respect to aid is proportional to the average marginal utility among college attendees. Consequently, it is lower for groups with higher parental transfers. In this setting, the average marginal utility governs both the marginal benefit and marginal cost of financial aid, allowing for a sharp characterization of the optimal policy.

Optimal Policy Across Groups. Assumption 1 simplifies the optimal policy formula (6):

Proposition 2 (Optimal Policy Formula)**.** *Suppose Assumption 1 holds. If $\hat{g}_h > 0$ and $p(b, \hat{g}_h) < 1$ for all b, then $\hat{g}_h$ satisfies*

$$\hat{g}_h = \sigma \left\{ \frac{1}{1+\lambda} - \frac{1}{\int V_1'(b + \hat{g}_h) m_{h|1}(b, \hat{g}_h) db} \right\}. \tag{8}$$

This formula shows that the optimal aid level increases with the marginal benefit of aid, which is captured by the average marginal utility of college attendees. To ensure that this marginal benefit declines with the aid level, I impose the following regularity condition on the distribution of parental transfers:

Assumption 2 (Log-Concave Distribution)**.** *$m_h(b)$ is log-concave in b for all h.*

Lemma 7 (Diminishing Marginal Benefit)**.** *Suppose that Assumptions 1 and 2 hold. If $p(b, g) < 1$ for all b, then $\int V_1'(b + g) m_{h|1}(b, g) db$ is decreasing in g.*

As financial aid increases, it directly raises each student's available resources, thereby reducing their marginal utility. At the same time, it draws in additional students with lower parental transfers, shifting the conditional distribution downward and pushing up the average marginal utility. Under log-concavity, the direct effect dominates, ensuring that the average marginal utility—and thus the right-hand side of (8)—declines with the level of aid.[21]

[21] Log-concavity is a property shared by many common distributions, including the normal, uniform, exponential, gamma, and beta distributions.

Proposition 3 (Optimal Policy Across Groups)**.** *Suppose Assumptions 1 and 2 hold, and that $p(b, g) < 1$ for all b and all g between $\hat{g}_h$ and $\hat{g}_{h'}$. (i) If $m_{h'}(b)/m_h(b)$ is increasing in b for all b, then $\hat{g}_{h'} \leq \hat{g}_h$; (ii) If $m_h(b)/m_{h'}(b)$ is unimodal in b, $\int b\left[m_{h|1}(b, g) - m_{h'|1}(b, g)\right]db = 0$ for some g between $\hat{g}_h$ and $\hat{g}_{h'}$, and $\gamma \geq 1/2$, then $\hat{g}_{h'} \geq \hat{g}_h$.*

This is the first main theoretical result. The optimal policy formula in (8) shows that differences in aid across groups are entirely driven by differences in the average marginal utility among college attendees. Lemmas 5 and 6 establish that average marginal utility is lower in groups with higher parental transfers or with less dispersion in the distribution of parental transfers. These results imply that the optimal financial aid amount decreases with the central tendency and increases with the dispersion of the parental transfer distribution within a group.

As documented in Section 2, both the central tendency and the dispersion of the parental transfer distribution increase with family resources. Therefore, part (i) of Proposition 3 implies that the optimal financial aid schedule is progressive: students from higher-income families, who tend to receive larger parental transfers, are awarded less aid—mirroring the current system. However, part (ii) shows that this progressivity is tempered by the greater inequality in parental transfers within high-income groups, which increases the average marginal utility of college attendees in those groups and calls for more generous aid.

The dispersion effect could, in principle, lead to a regressive aid schedule, and I examine its quantitative impact in Section 5. There is, however, a simple case where this effect is small:

Assumption 3 (Exponential Distribution)**.** *$m_h(b) = (e^{-b/\theta_h})/\theta_h$, where $\theta_h > 0$.*

Corollary 2 (Optimal Policy with Exponential Distribution)**.** *Suppose Assumptions 1 and 3 hold, and that $p(b, g) < 1$ for all b and all g between $\hat{g}_h$ and $\hat{g}_{h'}$. If $\theta_{h'} \geq \theta_h$, then $\hat{g}_{h'} \leq \hat{g}_h$.*

Under Assumption 3, parental transfers follow an exponential distribution. This distribution is log-concave (satisfying Assumption 2) and has both mean and standard deviation equal to θ_h. As a result, groups with higher θ_h exhibit both higher average transfers and greater dispersion. Nevertheless, the likelihood ratio $m_{h'}(b)/m_h(b)$ remains monotonic. By Proposition 3(i), it follows that groups with higher average transfers receive less financial aid.

3.4 Multiple Schooling Choices

Now, I extend the analysis to the case of multiple schooling options—such as two-year versus four-year colleges, as considered in the quantitative model of Section 4—that differ in both the cost and return to education. I characterize how the optimal amount of financial aid varies not only across parental income groups but also across different types of schooling choices.

Intensive Margin. Let $\mathcal{J}$ denote the set of college options, excluding the indices for non-attendance and attendance, denoted by 0 and 1. For each $j \in \mathcal{J}$, let y_j, k_j, and g_j represent the associated lifetime earnings, monetary cost, and financial aid, respectively. Higher levels of schooling involve greater monetary costs but also yield higher returns, so that $y_j \leq y_{j'}$ and $k_j \leq k_{j'}$ for $j' > j$. I also assume $k_j < \bar{l}$ for all $j \in \mathcal{J}$, ensuring that all college options remain feasible regardless of parental transfers or financial aid.

Let $V_j(b + g_j)$ denote the indirect utility from choosing option $j \in \mathcal{J}$. This function is analogous to the earlier definition of $V_1(b + g)$, but with option-specific values for lifetime earnings, cost, and aid—that is, y_j, k_j, and g_j, respectively. Individuals differ in their preferences over college options—or psychic returns—which are modeled as additive utility components ε_j. Given attendance, each individual sovles the following problem:

$$\max_{j \in \mathcal{J}} \left\{ V_j(b + g_j) + \varepsilon_j \right\}.$$

Let $\boldsymbol{g} := (g_j)_{j \in \mathcal{J}}$ denote the vector of financial aid amounts across options. Define $p_{j|1}(b, \boldsymbol{g})$ as the probability that an individual with parental transfer b chooses option j, conditional on attending college and given aid vector $\boldsymbol{g}$. Let $V_1(b, \boldsymbol{g})$ denote the corresponding expected utility from college attendance.[22]

Extensive Margin. The extensive margin of college choice remains unchanged: an individual enrolls in college if and only if $\varepsilon \leq V_1(b, \boldsymbol{g}) - V_0(b)$, where ε is independent of ε_j for all $j \in \mathcal{J}$. Let $p(b, \boldsymbol{g}) := F\big(V_1(b, \boldsymbol{g}) - V_0(b)\big)$ denote the overall college attendance rate, and define $p_j(b, \boldsymbol{g}) := p(b, \boldsymbol{g}) p_{j|1}(b, \boldsymbol{g})$ as the fraction of individuals who choose option j. Let $m_{h|j}(b, \boldsymbol{g}) := p_j(b, \boldsymbol{g}) m_h(b) / \int p_j(b', \boldsymbol{g}) m_h(b') db'$ denote the conditional density of parental transfers among individuals in group h who choose schooling option j.

Optimal Policy Across Groups and Schooling Levels. The social planner can only observe individuals' schooling choices and group membership, and sets $\boldsymbol{g}_h := (g_{j,h})_{j \in \mathcal{J}}$ for all $h \in \mathcal{H}$, where $g_{j,h}$ is the amount of financial aid for option j in group h. The efficiency losses and planning problem with multiple schooling choice are defined similarly as the case with binary schooling choice. Let $(\hat{\boldsymbol{g}}_h)_{h \in \mathcal{H}} := (\hat{g}_{j,h})_{(j,h) \in \mathcal{J} \times \mathcal{H}}$ denote the optimal policy.

For analytical characterization, I assume that intensive-margin psychic returns are independent across schooling options and impose a specific distribution on their marginal distributions.

Assumption 4 (Psychic Returns)**.** *For all $j \in \mathcal{J}$, $\varepsilon_j = \mu_j + \sigma_1 (z_j - E[z_j]) / SD(z_j)$, where z_j follows a distribution with cumulative distribution function* $\exp(-e^{-z_j})$*, and $\sigma_1 := \sigma\, SD(z_j)$.*

Under Assumption 4, each ε_j is a location-scale transformation of a standard Gumbel (Type I extreme value) distribution. This yields closed-form expressions for both $p_{j|1}(b, \boldsymbol{g})$

[22] Formal definitions, along with additional analytical details for this subsection, are provided in Section B.14 of the Appendix.

and $V_1(b, \boldsymbol{g})$. Together with Assumption 1, these distributional assumptions lead to a tractable optimal policy formula, which is analogous to (8):

$$\hat{g}_{j,h} = \sigma \left\{ \frac{1}{1+\lambda} - \frac{1}{\int V_j'(b + \hat{g}_{j,h}) m_{h|j}(b, \hat{\boldsymbol{g}}_h) db} \right\}. \tag{9}$$

This formula makes it clear that the previous result regarding the impact of parental transfer distributions (Proposition 3) extends to the case of multiple schooling choices.

Proposition 4 (Optimal Policy Across Groups by Schooling Level)**.** *Suppose Assumptions 1, 2, and 4 hold, and that $p(b, \boldsymbol{g}) < 1$ for all b and all $\boldsymbol{g}$ between $\hat{\boldsymbol{g}}_h$ and $\hat{\boldsymbol{g}}_{h'}$. (i) If $m_{h'}(b)/m_h(b)$ is increasing in b for all b, then $\hat{g}_{j,h'} \leq \hat{g}_{j,h}$ for all $j \in \mathcal{J}$; (ii) If $m_h(b)/m_{h'}(b)$ is unimodal in b, $\int b\left[m_{h|1}(b, \boldsymbol{g}) - m_{h'|1}(b, \boldsymbol{g})\right] db = 0$ for some $\boldsymbol{g}$ between $\hat{\boldsymbol{g}}_h$ and $\hat{\boldsymbol{g}}_{h'}$, and $\gamma \geq 1/2$, then $\hat{g}_{j,h'} \geq \hat{g}_{j,h}$ for all $j \in \mathcal{J}$.*

The formula (9) also implies that variation in optimal financial aid across schooling levels within group h depends only on differences in the average marginal utility across those levels. The following lemma identifies two opposing forces that shape this variation:

Lemma 8 (Marginal Utility Across Schooling Levels)**.** *Suppose that Assumptions 1 and 4 hold. For $j' > j$ with $g_j = g_{j'}$, (i) $V_{j'}'(b + g_{j'}) \geq V_j'(b + g_j)$ for all b, and (ii) $m_{h|j'}(b, \boldsymbol{g})/m_{h|j}(b, \boldsymbol{g})$ is increasing in b.*

On the one hand, individuals are more likely to be borrowing constrained at higher schooling levels, which involve both greater costs and returns. This raises their marginal utility for a given level of parental transfers. On the other hand, higher schooling levels tend to be selected by individuals with larger parental transfers, lowering the average marginal utility at those levels. If everyone in a group received the same amount of parental transfers, the first effect would dominate, implying that optimal financial aid increases with schooling level—a pattern consistent with the current financial aid system. However, when there is within-group inequality in parental transfers, the second effect may dominate, potentially reversing this pattern.

Proposition 5 (Optimal Policy Across Groups and Schooling Levels)**.** *Suppose Assumptions 1, 3, and 4 hold, and that $p(b, \boldsymbol{g}) < 1$ for all b and all $\boldsymbol{g}$ between $\hat{\boldsymbol{g}}_h$ and $\hat{\boldsymbol{g}}_{h'}$. (i) $\hat{g}_{j,h'} \leq \hat{g}_{j,h}$ for all j if $\theta_{h'} \geq \theta_h$; (ii) $\hat{g}_{j,h}$ increases in j if $\theta_h \leq \sigma / \max_{j \in \mathcal{J}} V_j'(0)$.*

Under Assumption 3, parental transfers follow an exponential distribution with both mean and standard deviation equal to θ_h. Interpreting higher θ_h as representing higher parental income, the optimal financial aid remains progressive across all schooling levels, consistent with Corollary 2. However, the degree of progressivity varies by schooling level and tends to be weaker at lower levels. For lower-income groups ($\theta_{h'} \leq \sigma / \max_{j \in \mathcal{J}} V_j'(0)$), the optimal policy allocates more aid to higher schooling levels than to lower ones. This reflects weaker negative selection into lower schooling due to limited within-group inequality in transfers. In contrast, for higher-income groups with greater inequality in transfers, stronger selection can reverse this pattern, leading to more aid at lower schooling levels. When this occurs, the aid gap between low- and high-income students is smaller at lower levels than at higher ones.[23]

4 Quantitative Model

This section presents and parameterizes a dynastic life-cycle model to examine the quantitative implications of unobserved heterogeneity in parental transfers for the design of financial aid. I begin by extending the framework to a multi-period setting that incorporates heterogeneity in ability and endogenous parental transfers. I then calibrate and estimate the model using data on the joint distribution of schooling, ability, parental income, and parental transfers.

4.1 A Dynastic Life-Cycle Model

Time is discrete, with each model period corresponding to a calendar year. A family consists of a parent and a child. At period $t = 0$, the child has completed high school and is about to enter college. Children live through period $t = T_k$ and parents live through period $t = T_p$.

Intertemporal Consumption Allocation. Children—refereed to as "youth"—derive utility from consumption c_t in each period t, and discount future utility at a subjective discount factor $\beta \in (0, 1)$. The present discounted value of lifetime money-metric utility (as of period $t = 0$) from a consumption path $\{c_t\}_{t=1}^{T_k}$ is given by:

$$U\left(\{c_t\}_{t=1}^{T_k}\right) := \Lambda \cdot \left(\sum_{t=1}^{T_k} \frac{\beta^t}{\Lambda} c_t^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}}, \text{ where } \Lambda := \sum_{t=1}^{T_k} \beta^t. \tag{10}$$

Let $j = 0$ denote the choice of not attending college, in which case youth begin working in period $t = 1$. Conditional on attending college, the set of schooling options is $\mathcal{J} := \{2, 3, 4\}$, where $j = 2$ corresponds to attending a two-year institution only, $j = 3$ to starting at a two-year institution and transferring to a four-year institution, and $j = 4$ to attending a four-year institution only. Let $\overline{\mathcal{J}} := \{0, 2, 3, 4\}$ be the full choice set, including the no-college option. Individuals attending only a two-year college (i.e., $j = 2$) study for $T_j = 2$ years, while those attending a four-year college (i.e., $j = 3$ or $j = 4$) study for $T_j = 4$ years.

Individuals can save and borrow at an annual gross interest rate $R = 1/\beta$. Let $y_j(a)$ denote the present discounted value of lifetime earnings for those with schooling j and ability a. Then,

[23] Suppose that $\hat{g}_{j',h} - \hat{g}_{j,h} \geq 0 \geq \hat{g}_{j',h'} - \hat{g}_{j,h'}$ for $j' > j$ and $h' > h$. This implies $\hat{g}_{j',h} - \hat{g}_{j',h'} \geq \hat{g}_{j,h} - \hat{g}_{j,h'}$.

the lifetime budget constraint for college attendees choosing $j \in \mathcal{J}$ is

$$\sum_{t=1}^{T_k} R^{-t} c_t \leq y_j(a) + b_j + g_j - k_j, \tag{11}$$

where k_j, b_j, and g_j are the present discounted values of college costs, parental transfers, and financial aid, respectively. The budget constraint for non-attendees is identical to (11), except that k_j and g_j are not included.

While individuals not enrolled in college can fully smooth consumption over time, those who are enrolled face borrowing constraints. For $j \in \mathcal{J}$,

$$\sum_{t=1}^{T_j} R^{-t} c_t + k_j - b_j - g_j \leq \bar{l}_j, \tag{12}$$

where $\bar{l}_j$ denotes the borrowing limit during college. These schooling-specific borrowing limits reflect a key feature of the U.S. federal student loan program, which offers higher annual and total loan limits to students in advanced years.

Let $V_0(b_0; a)$ denote the money-metric indirect utility of not attending college, conditional on (b_0, a), and let $V_j(b_j + g_j; a)$ represent the money-metric indirect utility of college choice $j \in \mathcal{J}$, conditional on (b_j, g_j, a). I define $V_j(b_j + g_j; a) := -\infty$ whenever $b_j + g_j < k_j - \bar{l}_j$, since in this case consumption cannot be positive, making choice j infeasible. Let $V_j'(b_j + g_j; a)$ be the derivative of $V_j(b_j + g_j; a)$ with respect to $b_j + g_j$. Then, a version of Lemma 1 holds:

Lemma 9. *For all $j \in \mathcal{J}$: (i) if the borrowing constraint (12) binds, then $V_j'(b_j + g_j; a) > 1$ and $V_j'(b_j + g_j; a)$ is strictly increasing in a; (ii) otherwise, $V_j'(b_j + g_j; a) = 1$.*

As discussed earlier, an additional dollar of initial wealth increases welfare by more than one dollar for borrowing-constrained individuals, as it enhances their ability to smooth consumption over time. This effect is particularly strong for higher-ability individuals who have higher post-schooling earnings and thus a greater desired level of consumption during schooling.

Education Choice. The nested choice structure described in Section 3.4 is maintained. First, I modify Assumption 1 so that the psychic cost of college, ε, is normalized to have mean zero and is assumed to have full support: specifically, $\mu = 0$ and z follows a standard logistic distribution with cumulative distribution function $1/(1 + e^{-z})$. Next, Assumption 4 is modified to capture heterogeneity in average psychic returns by ability a and parental income group h. For all $j \in \mathcal{J}$, the psychic return of college option j is specified as $\varepsilon_j = \mu_j(a, h) + \sigma_1(z_j - \mathrm{E}[z_j])/\mathrm{SD}(z_j)$, where z_j still follows a standard Gumbel distribution.

Under these assumptions, the model yields closed-form expressions for the choice probabilitie $p_{j|1}(\boldsymbol{b}, \boldsymbol{g}, a, h)$ and $p(\boldsymbol{b}, \boldsymbol{g}, a, h)$, as well as for the value functions $V_1(\boldsymbol{b}, \boldsymbol{g}, a, h)$ and

$V(\boldsymbol{b}, \boldsymbol{g}, a, h)$, where $\boldsymbol{b} := (b_j)_{j \in \overline{\mathcal{J}}}$ denotes the parental transfer schedule.[24] Furthermore, define $p_j(\boldsymbol{b}, \boldsymbol{g}, a, h) := p(\boldsymbol{b}, \boldsymbol{g}, a, h)\, p_{j|1}(\boldsymbol{b}, \boldsymbol{g}, a, h)$ and $p_0(\boldsymbol{b}, \boldsymbol{g}, a, h) := 1 - p(\boldsymbol{b}, \boldsymbol{g}, a, h)$.

Let $\tilde{\sigma} := \sigma/\mathrm{SD}(z)$ and $\tilde{\sigma}_1 := \sigma_1/\mathrm{SD}(z_j)$ denote the relative scale parameters, where $\mathrm{SD}(z) = \pi/\sqrt{3}$ and $\mathrm{SD}(z_j) = \pi/\sqrt{6}$. I impose the restriction $\tilde{\sigma} \geq \tilde{\sigma}_1$, which ensures that the correlation in psychic returns across college options, $\mathrm{Corr}(\varepsilon_j, \varepsilon_{j'}) = 1 - (\tilde{\sigma}_1/\tilde{\sigma})^2$, is nonnegative. A special case is $\tilde{\sigma} = \tilde{\sigma}_1$, corresponding to a standard multinomial logit model rather than a nested logit structure.

Parental Transfer Decision. At time $t = 0$, parents are endowed with wealth w, which represents the sum of initial net worth and the present discounted value of future earnings. Because parents are altruistic toward their child, they allocate their wealth between their own consumption and transfers to the child. The degree of altruism varies across families and is captured by a parameter $\delta \in [0, 1]$.

To simplify their decision problem, I assume that parents do not observe their child's psychic costs of college at the time of transfer, though they share all other information available to the child. Parents offer and commit to the transfer schedule $\boldsymbol{b}$. Then, a schooling option j is chosen with probability $p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)$, leading to an expected parental transfer of $\sum_{j \in \overline{\mathcal{J}}} p_j(\boldsymbol{b}, \boldsymbol{g}, a, h) b_j$. Therefore, the parent's expected annual consumption, c_p, satisfies

$$\sum_{t=1}^{T_p} R^{-t} c_p = w - \sum_{j \in \overline{\mathcal{J}}} p_j(\boldsymbol{b}, \boldsymbol{g}, a, h) b_j. \tag{13}$$

Similarly, the expected money-metric indirect utility of the child, $V(\boldsymbol{b}, \boldsymbol{g}, a, h)$, can be expressed in terms of annual consumption c_k satisfying

$$\sum_{t=1}^{T_k} R^{-t} c_k = V(\boldsymbol{b}, \boldsymbol{g}, a, h). \tag{14}$$

Although parents face uncertainty about their child's schooling preferences, they are risk neutral and care only about expected values, c_p and c_k. Their preferences are represented by

$$(1-\delta) \sum_{t=1}^{T_p} \beta^t v(c_p) + \delta \sum_{t=1}^{T_k} \beta^t v(c_k), \tag{15}$$

where $v(c) := (c^{1-1/\eta} - 1)/(1 - 1/\eta)$. The parameter $\eta \geq 0$ denotes the elasticity of intergenerational substitution (EGS), which captures the strength of the desire to smooth consumption across generations (Córdoba and Ripoll, 2018). It may differ from the EIS, γ, in which case the parental utility function in (15) is more general than the specification typically assumed in

[24] These expressions are provided in Section B.19 of the Appendix.

the literature, as in Becker and Tomes (1986). Allowing the EGS to differ from the EIS enables the model to better match the empirical relationship between educational attainment, ability, parental transfers, and parental income.

Taking $(p_j(\boldsymbol{b}, \boldsymbol{g}, a, h))_{j \in \overline{\mathcal{J}}}$ and $V(\boldsymbol{b}, \boldsymbol{g}, a, h)$ as given, the parent chooses $\boldsymbol{b}$ to maximize (15), subject to (13), (14), and the constraint $b_j \in [0, w]$ for all $j \in \overline{\mathcal{J}}$. The following proposition characterizes parental transfers:

Proposition 6. *Consider a family with $(\boldsymbol{g}, a, h, w, \delta)$, and suppose that $\hat{\boldsymbol{b}} := (\hat{b}_j)_{j \in \overline{\mathcal{J}}}$ solves the parent's problem, and that $\hat{b}_j + g_j \geq k_j - \bar{l}_j$ for all $j \in \overline{\mathcal{J}}$. Then: (i) $\hat{b}_j = \hat{b}_{j'}$ for all j and j' such that the constraint (12) does not bind; (ii) If $\tilde{\sigma} = \tilde{\sigma}_1$, then for any j with $\hat{b}_j \in (0, w)$,*

$$\hat{b}_j = \sum_{j' \in \overline{\mathcal{J}}} p_{j'}(\hat{\boldsymbol{b}}, \boldsymbol{g}, a, h)\, \hat{b}_{j'} + \tilde{\sigma} \left\{ \left(\frac{\delta}{1-\delta} \right) \frac{v'(\hat{c}_k)}{v'(\hat{c}_p)} - \frac{1}{V_j'(\hat{b}_j + g_j; a)} \right\}, \tag{16}$$

where $\hat{c}_p$ and $\hat{c}_k$ are given by (13) and (14), evaluated at $\boldsymbol{b} = \hat{\boldsymbol{b}}$.

Parental behavior is particularly transparent in the multinomial logit case, where condition (16) holds. When this condition applies to all j, taking expectations over j yields

$$(1-\delta) v'(\hat{c}_p) = \delta v'(\hat{c}_k) \left\{ \sum_{j \in \overline{\mathcal{J}}} \frac{p_j(\hat{\boldsymbol{b}}, \boldsymbol{g}, a, h)}{V_j'(\hat{b}_j + g_j; a)} \right\}^{-1}. \tag{17}$$

When all borrowing constraints are slack, condition (17) simplifies to $(1-\delta)v'(\hat{c}_p) = \delta v'(\hat{c}_k)$. In this case, wealthier or more altruistic parents provide larger transfers. In contrast, higher-ability youth receive smaller transfers, since they earn and consume more. How transfers vary with these factors depends on the EGS, while the EIS does not affect behavior in this case.

When the borrowing constraint binds for some j, (17) implies $(1-\delta)v'(\hat{c}_p) > \delta v'(\hat{c}_k)$. Now the EIS becomes relevant: with a stronger preference for intertemporal smoothing, consumption fluctuations are more costly. If the EIS is small relative to the EGS, higher-ability youth may receive larger parental transfers despite their higher consumption, because their consumption distortion is greater.

Condition (16) explains why non-paternalistic parents may condition transfers on their child's schooling choice.[25] Parental transfers differ across schooling levels only when borrowing constraints bind and parents lack full information about their child's preferences. Otherwise, non-paternalistic parents have no incentive to influence schooling choices through differential transfers.[26] With incomplete information, transfers are larger for schooling options where children's consumption is more distorted by borrowing constraints.

[25]Such "conditional parental transfer rules" are often assumed exogenously (Keane and Wolpin, 2001; Johnson, 2013) or motivated by paternalistic preferences (Abbott et al., 2019; Colas, Findeisen, and Sachs, 2021).

[26]Hence, dynamic models with complete information typically feature lump-sum parental transfers (e.g., Becker and Tomes, 1986; Krueger and Ludwig, 2016).

4.2 NLSY97 Data

I use data from the NLSY97, a longitudinal survey of 8,984 Americans born between 1980 and 1984. The survey was conducted annually from 1997 to 2011 and biennially thereafter. It provides rich information on respondents' educational choices and detailed data on family background. Crucially, it includes questions on parental transfers, which have been used in prior work to study their role in educational decisions (e.g., Johnson, 2013). Because data on parental transfers are available only through 2010—when the youngest respondents were age 26—I construct a measure of total transfers received between ages 18 and 26, discounted back to age 17. This serves as the primary measure of parental transfers in the analysis.

As noted by Abbott et al. (2019), constructing the parental transfer variable in the NLSY97 data involves judgment (see Section C.2 of the Appendix for details), which may introduce measurement error. Moreover, because transfers are observed only during early adulthood, they are not directly comparable to those in the model, which represent lifetime transfers. Individuals receiving substantial support during college are likely to receive additional transfers later, as parents may prefer to spread transfers over time rather than front-load them when children are young.[27] As a result, observed transfers likely understate total lifetime support, implying that measurement error is potentially non-classical.[28]

To address potential measurement error in reported parental transfers, I use the value of education savings accounts—collected as part of parental net worth—as a secondary measure. These tax-advantaged accounts are designed to help families save for future college expenses and can be interpreted as planned parental transfers, since parents cannot reclaim the funds for personal use without incurring penalties. As such, their value serves as an additional indicator of actual parental support. For instance, if parents report substantial holdings in these accounts but the child reports receiving no college support, the discrepancy may reflect a reporting error.

Schooling choices are classified using monthly full-time enrollment records from two- and four-year colleges. Individuals who attend only a two-year college are assigned to the two-year only group ($j = 2$), while those who begin at a two-year college and later transfer to a four-year institution are classified as choosing the two- and four-year college option ($j = 3$). Those whose first enrollment is at a four-year institution are assigned to the four-year only group ($j = 4$).

As a measure of ability, I use quartiles of Armed Forces Qualifying Test (AFQT) scores.[29] Parental background information—including income, net worth, education, and age at the time of the youth's birth—is taken from the first round of parent interviews. To improve precision, I also average parental income reports from rounds 2 through 5, primarily collected for youth

[27]For instance, parents may use bequests to compensate children for services (Bernheim, Shleifer, and Summers, 1985), or withhold transfers fearing that early generosity could lead to future demands (Bruce and Waldman, 1990).

[28]An alternative would be to modify the model to allow multiple transfers, aligning more closely with the data. However, as shown by Brown, Scholz, and Seshadri (2012), repeated interactions introduce inefficiencies—even in the absence of borrowing constraints—that complicate the policy problem.

[29]AFQT scores are widely used to measure cognitive ability. Most respondents took the test as part of the survey.

still living at home. When both parents are present, parental education refers to the highest completed grade of either parent, and parental age is the average of both parents' ages. Parental education is categorized into four groups: no college (≤12 years), some college (13–15 years), college (16 years), and more than college (≥17 years). The analysis also accounts for family composition—specifically, the number of parents and siblings present in the household in round one—as well as race, grouped into four mutually exclusive categories: white, Black, Hispanic (including other races), and Asian.

Solving the model and comparing it with the data requires several additional assumptions. First, the model assumes a simple family structure consisting of one parent and one child. To account for differences in family size observed in the data, I allow family composition to influence parental altruism in the model. Second, I assume that parents receive a constant stream of income starting from the year their child turns 17. Parental lifetime wealth, w, is then defined as the sum of their net worth and the present discounted value of remaining income.

I restrict the sample to youth from the random samples, excluding those from the minority and poor white oversamples. I include individuals who were under age 18 and living with their parents in the first survey round. I exclude those who first enrolled in college after age 20 or were still enrolled after age 26, as older students may be considered independent for financial aid purposes. I also drop individuals with missing data on any key variables or with nonpositive parental wealth.

4.3 Calibration

The parameterization proceeds in two stages. I begin by externally setting parameters related to the monetary returns and costs of schooling. The remaining preference parameters are then estimated through model simulation, as described in the next subsection. All monetary values in this and the following sections are expressed in 2004 U.S. dollars, adjusted using the Consumer Price Index (CPI-U-RS).

Monetary Returns to Schooling. Youth are age 17 in period $t = 0$. All individuals work until age 65 and live until age 80. The present discounted value of lifetime after-tax earnings for those choosing schooling option $j \in \overline{\mathcal{J}}$ is given by

$$y_j(a) = \sum_{x=1}^{65-17-T_j} R^{-(x+T_j)} \left[\tilde{y}_j(a,x) - \mathcal{T}\big(\tilde{y}_j(a,x)\big)\right],$$

where $\tilde{y}_j(a,x)$ denotes before-tax annual earnings for individuals with schooling j, ability a, and potential experience x, and $\mathcal{T}(\cdot)$ is the income tax function. I assume a constant annual interest rate of 3%, so that $R = 1/\beta = 1.03$. The tax function follows Guner, Kaygusuz, and Ventura (2014): $\mathcal{T}(\tilde{y}) = 0.264[1 - (0.012\tilde{y}^{0.964} + 1)^{-1/0.964}]\tilde{y}$.

The parameters of the earnings function are estimated using NLSY97 data from 1997 to 2019. I apply the same sample selection criteria as in Section 4.2, except that the restrictions based on family-related variables are omitted. To estimate annual earnings $\tilde{y}_j(a, x)$, I regress log real annual earnings—before-tax income from all jobs—on indicators for schooling choices, AFQT quartiles, and potential experience levels. I then assume that earnings remain constant after 16 years of experience when constructing the present discounted value of lifetime earnings. Section C.1 of the Appendix presents the Ordinary Least Squares (OLS) estimates from this regression (Table C.1) and the resulting present discounted values of lifetime earnings, $y_j(a)$ (Table C.2).

Monetary Costs of Schooling. Monetary costs and financial aid for each level of schooling are computed using data from the 2003–2004 National Post-Secondary Student Aid Survey (NPSAS:04), restricted to full-time, full-year, dependent students who applied for federal financial aid, did not live at home, and were not paying out-of-jurisdiction tuition. The NPSAS:04 is used because the 2003–2004 academic year overlaps with the period when most NLSY97 respondents were enrolled in college. The monetary cost k_j is defined as the present discounted value of the average annual cost of tuition and books, net of earnings while enrolled. Financial aid $g_{j,h}$ reflects the present discounted value of the average annual amount of grant aid. To capture key features of the need-based aid system in a tractable way, I divide the population into quartiles of parental income and estimate average grant aid separately for each group.[30] Given the limited role of private student loans in 2003–2004, I assume students borrow only from government sources. The total borrowing limit $\bar{l}_j$ is based on the annual limits of the Stafford Loan Program for that year. Tables C.3 and C.4 in Section C.1 of the Appendix report the estimated annual and total amounts.

For the two-year only option ($j = 2$) and the four-year only option ($j = 4$), total costs and aid are based directly on the annual amounts for their respective institution types. For the two- and four-year option ($j = 3$), total amounts are computed as a weighted sum of the annual figures for two- and four-year institutions. Accordingly, the grant aid for this option, $g_{3,h}$, is calculated as:

$$g_{3,h} = g_{2,h} + \left(\frac{\sum_{t=3}^{4} R^{-t}}{\sum_{t=1}^{4} R^{-t}} \right) g_{4,h}, \quad \forall h \in \mathcal{H}. \tag{18}$$

This relationship is maintained when solving for the optimal policy.

[30]Because the NPSAS:04 includes only college enrollees, its parental income distribution differs from that of the NLSY97. Therefore, I apply the parental income quartile thresholds from the NLSY97 to the NPSAS:04 sample.

4.4 Maximum Likelihood Estimation

Assuming parental transfers are measured with error, the remaining parameters to estimate include: the consumption smoothing parameters (γ, η); the distribution of psychic returns $\big(\mu_2(a,h), \mu_3(a,h), \mu_4(a,h), \sigma, \sigma_1\big)$; the distribution of parental altruism; and the distribution of measurement error in reported transfers. They are estimated using the NLSY97 data.

Specification. Let $\boldsymbol{X}$ denote the vector of predetermined variables—financial aid schedule, youth ability, parental income, wealth, and age—that influence parental transfer decisions, conditional on parental altruism. Let $\hat{\boldsymbol{b}}(\delta, \boldsymbol{X})$ be the solution to the parent's problem for a family with characteristics $\boldsymbol{X}$ and parental altruism level δ. The implied probability of choosing schooling option j, conditional on $(\delta, \boldsymbol{X})$, is given by $\Pr(J = j|\delta, \boldsymbol{X}) = p_j\big(\hat{\boldsymbol{b}}(\delta, \boldsymbol{X}), \boldsymbol{g}_h, a, h\big)$, where J is a random variable representing schooling attainment.

Let $\boldsymbol{Z}$ denote a set of additional categorical, predetermined variables with finite support $\mathcal{Z} \coloneqq \{z_q\}_{q=1}^{Q}$. The vector $\boldsymbol{Z}$ consists of 9 indicator variables: 3 each for parental education, race, and family composition. These variables are excluded from $\boldsymbol{X}$ but may be correlated with parental altruism δ. Altruism is assumed to take values on a finite support $\mathcal{D} \coloneqq \{\delta_m\}_{m=1}^{M} = \{0.1, 0.3, 0.5, 0.7\}$, with its distribution depending on $\boldsymbol{Z}$ and independent of $\boldsymbol{X}$ given $\boldsymbol{Z}$:

$$\Pr(\delta = \delta_m|\boldsymbol{X}, \boldsymbol{Z}) = \Pr(\delta = \delta_m|\boldsymbol{Z}) = \Pr\left(\boldsymbol{Z}^\top \boldsymbol{\alpha} + \upsilon \in (\chi_{m-1}, \chi_m]|\boldsymbol{Z}\right),$$

where $\chi_0 \coloneqq -\infty$, $\chi_M \coloneqq \infty$, and υ is a standard logistic random variable. Note that $\boldsymbol{Z}$ exclude a constant term, as the intercept cannot be separately identified from the thresholds $\{\chi_m\}_{m=1}^{M-1}$.

Let $B^* \coloneqq \hat{b}_J(\delta, \boldsymbol{X})$ denote the unobserved true parental transfer. Let B and $\tilde{B}$ be two noisy observed measures: reported transfers and education savings account values, respectively. Conditional on B^*, these are assumed independent of each other and of other variables:

$$\Pr(B \in \mathcal{B}_k, \tilde{B} \in \tilde{\mathcal{B}}_l|B^*, \boldsymbol{X}, \boldsymbol{Z}, J) = \Pr(B \in \mathcal{B}_k|B^*)\Pr(\tilde{B} \in \tilde{\mathcal{B}}_l|B^*). \tag{19}$$

Here, $\{\mathcal{B}_k\}_{k=1}^{K} = \big\{\{0\}, (0,1], (1,5], (5,20], (20,50], (50,\infty)\big\}$ and $\{\tilde{\mathcal{B}}_l\}_{l=1}^{L} = \big\{\{0\}, (0,20], (20,\infty)\big\}$ partition $\mathbb{R}_+$ (in \$1,000). I assume $\Pr(B \in \mathcal{B}_k|B^*) = \Pr(B \in \mathcal{B}_k|B^* \in \mathcal{B}_{k^*})$ and $\Pr(\tilde{B} \in \tilde{\mathcal{B}}_l|B^*) = \Pr(\tilde{B} \in \tilde{\mathcal{B}}_l|B^* \in \mathcal{B}_{k^*})$ for $B^* \in \mathcal{B}_{k^*}$, and parameterize these probabilities using multinomial logits:

$$\Pr(B \in \mathcal{B}_k|B^* \in \mathcal{B}_{k^*}) = \frac{\exp(\zeta_{k,k^*})}{\sum_{k'=1}^{K}\exp(\zeta_{k',k^*})}, \quad \Pr(\tilde{B} \in \tilde{\mathcal{B}}_l|B^* \in \mathcal{B}_{k^*}) = \frac{\exp(\xi_{l,k^*})}{\sum_{l'=1}^{L}\exp(\xi_{l',k^*})},$$

with $\zeta_{K,k^*} = \xi_{L,k^*} = 0$ for all k^*. To ensure identification, I impose that $\Pr(B \in \mathcal{B}_1|B^* \in \mathcal{B}_{k^*})$ strictly decreases in k^*, and analogously for $\tilde{B}$. This structure flexibly captures non-classical measurement error while remaining tractable.

Finally, I specify the average psychic returns additively as $\mu_j(a,h) = \psi_j + \kappa_a + \nu_h$, with the normalization $\kappa_a = \nu_h = 0$ for some (a, h).

Likelihood. The joint probability of schooling and the two parental transfer measures is

$$\Pr(J = j, B \in \mathcal{B}_k, \tilde{B} \in \tilde{\mathcal{B}}_l | \boldsymbol{X}, \boldsymbol{Z}) = \sum_{m=1}^{M} \Pr(\delta = \delta_m | \boldsymbol{Z}) \Pr(J = j | \delta = \delta_m, \boldsymbol{X})$$
$$\times \left[\sum_{k^*=1}^{K} \mathbb{I}_{\hat{b}_j(\delta_m, \boldsymbol{X}) \in \mathcal{B}_{k^*}} \Pr(B \in \mathcal{B}_k | B^* \in \mathcal{B}_{k^*}) \Pr(\tilde{B} \in \tilde{\mathcal{B}}_l | B^* \in \mathcal{B}_{k^*}) \right], \tag{20}$$

using the conditional independence assumptions $\Pr(J|\delta, \boldsymbol{X}, \boldsymbol{Z}) = \Pr(J|\delta, \boldsymbol{X})$ and $\Pr(\delta|\boldsymbol{X}, \boldsymbol{Z}) = \Pr(\delta|\boldsymbol{Z})$, along with (19). Based on (20), I construct the likelihood of an observation $(J, B, \tilde{B}, \boldsymbol{X}, \boldsymbol{Z})$. Parameter identification is discussed in Section C.3 of the Appendix.

Parameter Estimates. Table 1 reports the estimated preference parameters. Panel A shows that the EIS is less than 1, consistent with most estimates surveyed by Browning, Hansen, and Heckman (1999). In contrast, the EGS exceeds 1, aligning with the findings of Córdoba and Ripoll (2018). Since the EGS is greater than the EIS, the model implies a weaker preference for smoothing consumption across generations than within a generation (i.e., over time). As discussed below, this feature helps the model replicate the observed positive relationship between parental transfers and youth ability.

The large negative estimates of ψ_j reported in Panel B of Table 1 indicate substantial psychic costs associated with attending college, consistent with prior findings surveyed by Heckman, Lochner, and Todd (2006). Attending any type of college entails a psychic cost exceeding \$160,000 in present value terms. However, these costs are smaller for youth with higher parental income or greater ability. Psychic returns also exhibit substantial heterogeneity, even after conditioning on parental income and ability, as reflected in the large estimated standard deviation. Although the model allows psychic returns across college types to be positively correlated, the estimated values of σ and σ_1 imply a correlation of only about 0.1.

Panel C of Table 1 presents the estimated parameters for the latent parental altruism index, $\boldsymbol{\alpha}$, along with the thresholds that define the ranges corresponding to each altruism type δ_m. Parental altruism is estimated to be higher among more educated parents and among racial minorities.[31] Family composition also affects altruism, though its influence is smaller compared to that of parental education.

Figure 2 shows the conditional probabilities of observed parental transfers given the true transfer amount, $\Pr(B \in \mathcal{B}_k | B^* \in \mathcal{B}_{k^*})$.[32] While higher actual transfers generally correspond to higher reported transfers, there is substantial dispersion due to measurement error. Reported

[31]Figure C.1 in the Appendix visualizes the resulting distribution of altruism across groups.

[32]The corresponding distribution for the secondary measure of parental transfers is shown in Figure C.2 in the Appendix.

Table 1: Estimated Preference Parameters

Parameter	Description	Estimate	
	A. Consumption Smoothing		
γ	Elasticity of intertemporal substitution	0.278*	(0.004)
η	Elasticity of intergenerational substitution	3.613*	(0.086)
	B. Psychic Returns to Schooling ($1,000s)		
ψ_2		-183.400*	(2.206)
ψ_3	Mean components for schooling choices	-203.550*	(2.032)
ψ_4		-162.800*	(1.870)
ν_2		10.405*	(1.663)
ν_3	Mean components for parental income quartiles	21.815*	(2.096)
ν_4		23.349*	(3.511)
κ_2		11.407*	(1.793)
κ_3	Mean components for ability quartiles	40.739*	(3.913)
κ_4		72.090*	(3.082)
σ	Standard deviation for the extensive margin	71.192*	(0.881)
σ_1	Standard deviation for the intensive margin	47.763*	(0.867)
	C. Determinants of Parental Altruism		
Some College		0.760*	(0.110)
College	Effects of parental education	1.787*	(0.171)
> College		2.270*	(0.200)
Black		0.819*	(0.202)
Hispanic	Effects of race	0.215	(0.157)
Asian		1.887	(1.278)
Two Parents		0.040	(0.050)
1 Sibling	Effects of family composition	0.473*	(0.113)
2+ Siblings		-0.039*	(0.016)
χ_1		0.136*	(0.062)
χ_2	Thresholds for parental altruism types δ_m	2.716*	(0.121)
χ_3		5.478	(13.955)

Notes: Standard errors, shown in parentheses, are calculated from 100 bootstrap replications.
* indicates statistical significance at the 5% level.

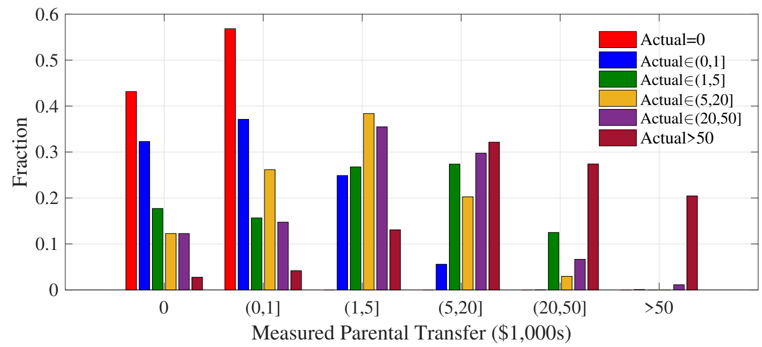


Figure 2: Distribution of Measured Parental Transfers Conditional on Actual Transfer Amounts

transfers tend to be biased downward, especially at higher levels of true transfers. As noted earlier, this downward bias may reflect that transfers are measured only during early adulthood, potentially omitting a significant portion of lifetime transfers.

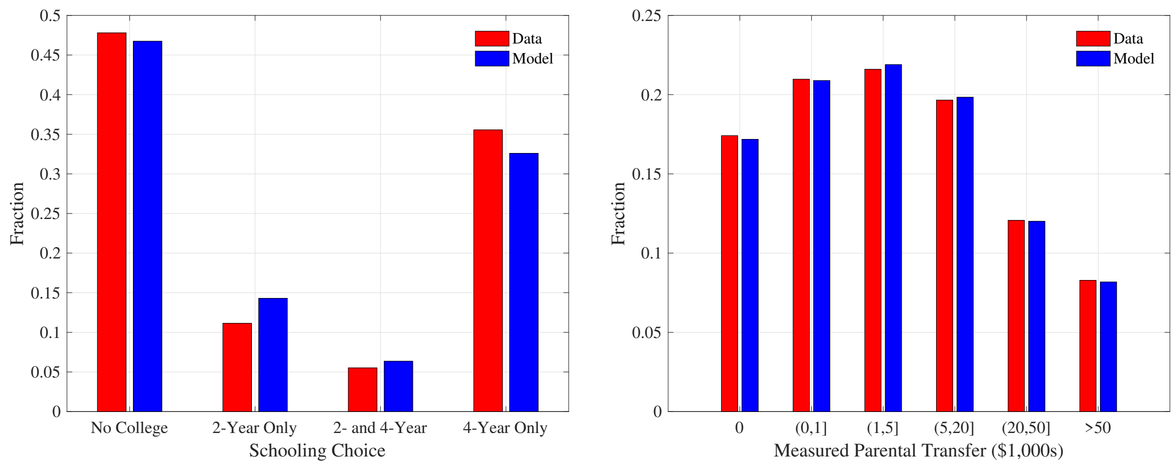
Model Fit. I now assess the fit of the estimated model with respect to two key outcomes: schooling choice and parental transfers. Figure 3 shows that the model replicates the marginal distributions of these outcomes reasonably well, although it overpredicts attendance at two-year colleges. To evaluate how well the model captures the joint distribution of outcomes and observed characteristics, I compare the estimated effects of covariates in the model to those in the data. Specifically, Figure 4 presents OLS estimates of how observed characteristics are associated with college attendance and with reported parental transfers exceeding $5,000.

As shown in prior research (e.g., Belley and Lochner, 2007), college attendance is positively associated with both parental income and youth ability. The model reproduces this pattern reasonably well, though the estimated effects are somewhat attenuated. In the model, the positive relationship between parental income and attendance reflects larger parental transfers interacting with borrowing constraints, as well as lower psychic costs of college. This occurs despite higher-income students receiving less need-based financial aid.

In the model, higher-ability youth are also more likely to attend college, driven by higher monetary returns and lower psychic costs, as well as greater parental transfers. This pattern persists despite a strong preference for intertemporal consumption smoothing, which could discourage attendance by amplifying the welfare cost of low consumption during college—especially for high-ability youth, for whom borrowing constraints are more likely to bind.

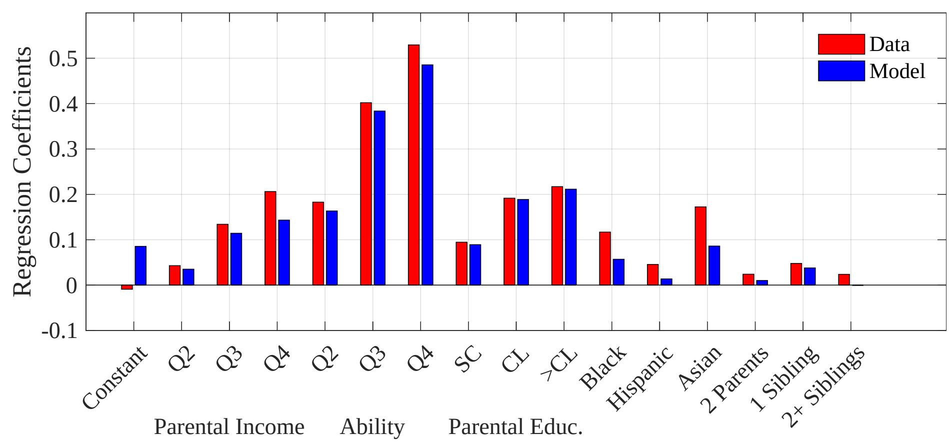
The estimated effect of ability on parental transfers is weaker than its effect on college attendance, likely reflecting intergenerational consumption smoothing—parents may give less to higher-ability children due to their greater lifetime income. Still, under weaker intergenerational than intertemporal smoothing, borrowing constraints induce a positive relationship between ability and transfers, consistent with the data.

The remaining covariates influence both college attendance and parental transfers in the

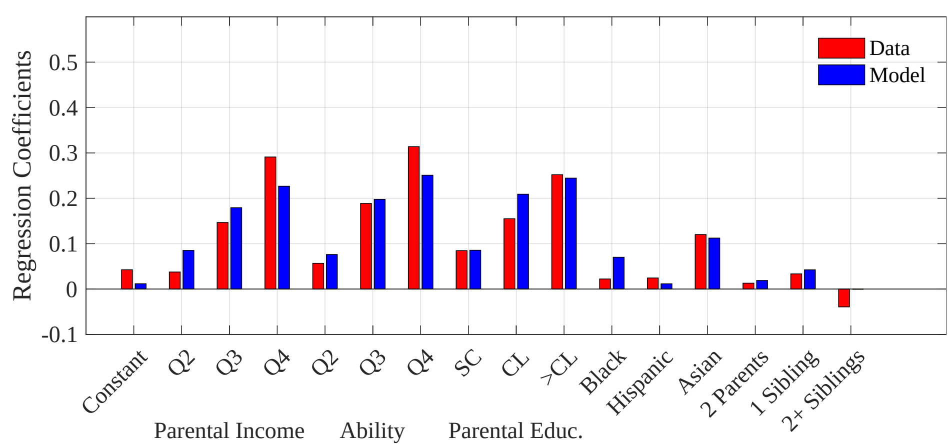


(a) Schooling Choice (b) Parental Transfer

Figure 3: Model Fit: Schooling Choice and Parental Transfer Distributions



(a) Effects on College Attendance



(b) Effects on Measured Parental Transfers $5,000+

Figure 4: Model Fit: Estimated Effects of Observed Characteristics on Key Outcomes

model through their effect on parental altruism. As discussed earlier, higher parental education and racial minority status are associated with greater altruism, leading to more generous transfers and higher college attendance. While the overall empirical patterns align with the model, the effects of being Black or Asian on college attendance are substantially larger than their effects on parental transfers—but only in the data.[33]

5 Optimal Policy Simulations

This section uses the estimated model to calculate optimal policies. I begin with a simplified scenario featuring exogenous parental transfers, consistent with the theoretical framework presented in Section 3. I then extend the analysis to the case of endogenous parental transfers.

5.1 Planning Problem

As discussed in Section 2, the amount of aid in the U.S. depends on a diverse set of factors, including parental income, assets, family composition, and parental age. However, solving for aid schedules that incorporate the full range of this information is computationally challenging. Instead, I approximate the current financial aid system by assigning distinct aid amounts to each parental income quartile, while removing within-quartile income variation and all differences in family composition and parental age. Specifically, I assign the average parental income within each quartile to all individuals in that quartile, assume all families consist of one parent and one child—removing heterogeneity in parental altruism due to family structure—and hold parental age constant across families. I retain variation in parental net worth within income quartiles, however, since it plays a relatively minor role under the current financial aid system.

I continue to assume that the social planner observes only each individual's schooling choice and parental income group, and sets aid amounts $(\boldsymbol{g}_h)_{h\in\mathcal{H}}$ to minimize aggregate inefficiencies. Let $i = 1, \ldots, N$ index individuals in the NLSY97 data. The planner's problem is:

$$\begin{aligned} \min_{(\boldsymbol{g}_h)_{h\in\mathcal{H}}} \frac{1}{N}\sum_{i=1}^{N}\sum_{m=1}^{M} \Pr(\delta=\delta_m|\boldsymbol{Z}_i) L\big(\hat{\boldsymbol{b}}(\delta_m,\boldsymbol{X}_i),\boldsymbol{g}_{h_i},a_i,h_i\big) \\ \text{subject to } \frac{1}{N}\sum_{i=1}^{N}\sum_{m=1}^{M} \Pr(\delta=\delta_m|\boldsymbol{Z}_i)\sum_{j\in\mathcal{J}} p_j\big(\hat{\boldsymbol{b}}(\delta_m,\boldsymbol{X}_i),\boldsymbol{g}_{h_i},a_i,h_i\big) g_{j,h_i} \le G, \end{aligned} \tag{21}$$

together with the constraints $g_{j,h} \ge 0$ for all $(j,h) \in \mathcal{J}\times\mathcal{H}$ and constraint (18), where $\hat{\boldsymbol{b}}(\delta,\boldsymbol{X})$ solves the parent's problem for a family with $(\delta,\boldsymbol{X})$.[34]

[33] This discrepancy may reflect omitted mechanisms, such as race being correlated with returns to college.

[34] See Sections D.1 and D.2 in the Appendix for the definition of inefficiencies in the quantitative model and the computational algorithm used to solve the planner's problem.

5.2 Exogenous Parental Transfers and Binary Schooling Choice

I begin with a simplified setup that closely mirrors the theoretical framework presented in Section 3.3. First, I assume that parental transfers are lump-sum and exogenous, which facilitates quantifying the impact of the transfer distribution on the optimal financial aid schedule. Specifically, for each individual, I compute a fixed lump-sum parental transfer—i.e., $\hat{b}_j(\delta, \boldsymbol{X}) = \hat{b}_{j'}(\delta, \boldsymbol{X})$ for all (j, j')—that solves the parent's problem under the current aid policy. These transfer amounts are then held fixed when evaluating alternative aid schedules. Second, I restrict the schooling choice to a binary decision: attend a four-year college or not attend college at all. Under this assumption, I characterize the optimal financial aid schedule for four-year colleges only.

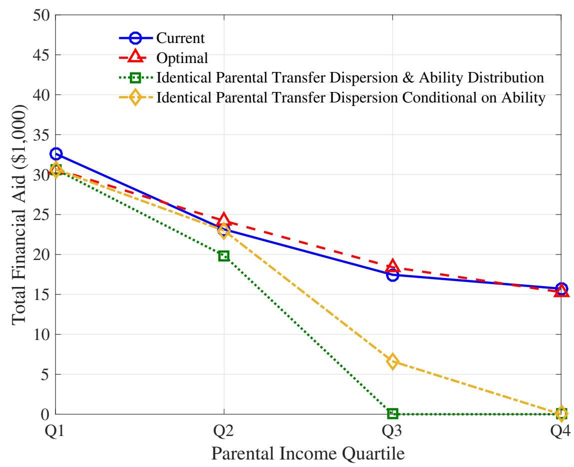


Figure 5: Financial Aid Schedules: 4-Year College, Exogenous Parental Transfers

Figure 5 compares the current (solid line) and budget-neutral optimal (dashed line) financial aid schedules for four-year colleges, both of which allocate approximately $7,500 per individual. Under the existing policy, students in the lowest parental income quartile receive about $33,000 in total financial aid over four years. Due to the progressivity of need-based aid, this amount declines to roughly $16,000 for students in the highest income quartile. Notably, the current aid schedule is already quite close to optimal.

What drives the degree of progressivity—or the slope—of the optimal financial aid schedule across parental income quartiles? In this framework, individuals differ only along two dimensions: parental transfers, and the monetary and psychic returns to college, with the latter shaped by both ability and parental income. Consequently, any progressivity in the optimal aid schedule must arise from how the joint distribution of these two factors varies across income quartiles. To isolate the contribution of each factor, I construct counterfactual distributions that hold specific features of the joint distribution fixed across quartiles, while allowing other features to vary. I then recalculate the optimal aid schedule under each counterfactual, holding the Lagrange multiplier on the constraint (21) fixed at its value under the actual distribution.

In the first counterfactual, I focus on the role of average parental transfers. I assume that individuals in the top three income quartiles are identical to those in the bottom quartile in all respects, except for a mean shift in their parental transfer distributions. Specifically, I construct the joint distribution of parental transfers and ability for each of the top three quartiles by taking the bottom quartile's joint distribution and adding the respective average parental transfer difference to every observation. This procedure preserves the inter-quartile differences in average transfers while keeping the dispersion of the transfer distribution and the distribution of ability identical across quartiles. To remove remaining variation in psychic returns by income, I also set $\nu_h = 0$ for $h \in \{2, 3, 4\}$.

The dotted line of Figure 5 shows the financial aid schedule under the first counterfactual distribution. The aid amount for the bottom quartile is unchanged from the actual optimal aid schedule (dashed line), since both the distribution for the bottom quartile and the Lagrange multiplier remain fixed. For the higher quartiles, however, aid amounts are lower under the counterfactual and drop to zero for the third and fourth quartiles. This pattern illustrates that, all else equal, an increase in average parental transfers reduces the optimal aid amount, consistent with Proposition 3(i). The fact that the counterfactual aid schedule is much steeper than the actual optimal aid schedule suggests that other inter-quartile differences—such as the greater dispersion of parental transfers and higher average ability for higher quartiles—mitigate the strong progressivity implied by average transfer differences alone.

In the second counterfactual, I incorporate inter-quartile differences in monetary and psychic returns to college, which were abstracted from in the first experiment. As in the first counterfactual, I start by assigning to each of the top three quartiles the joint distribution of parental transfers and ability from the bottom quartile. I then reweight individuals based on ability—following the approach of DiNardo, Fortin, and Lemieux (1996)—so that the marginal distribution of ability in each quartile matches that of the actual distribution. This ensures that, under the counterfactual, the dispersion of parental transfers conditional on ability is identical across quartiles. To preserve the observed inter-quartile differences in average parental transfers, I add a quartile-specific constant to the parental transfer of every individual in the top three quartiles. Finally, I account for variation in psychic returns across income quartiles by allowing the parameter ν_h to vary with h, as in the baseline model.

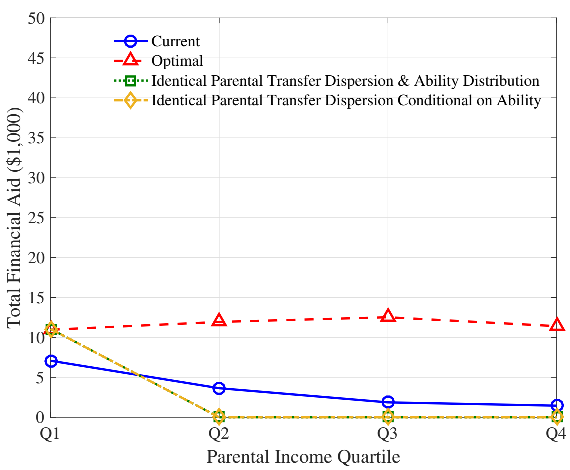
Figure 5 shows that the financial aid schedule under the second counterfactual distribution (dash-dotted line) is less progressive than that under the first counterfactual (dotted line). This reflects a positive correlation between ability and parental income. Lemma 9 implies that, for a given level of parental transfer, higher-ability individuals are more likely to be borrowing constrained and therefore benefit more from financial aid. While Figure 4(b) shows that individuals with higher ability or higher parental income tend to receive larger parental transfers, these transfers may not fully offset their greater need for resources to smooth consumption over time. As a result, incorporating higher average ability for higher parental income quartiles reduces the aid amounts for these groups less sharply, leading to a flatter aid schedule than in the first counterfactual.

Relative to the actual optimal aid schedule (dashed line), however, the aid schedule under the second counterfactual remains more progressive. This is because, in the actual distribution, parental transfers are more dispersed—conditional on ability—for higher-income quartiles, whereas the counterfactual imposes constant dispersion across quartiles. Proposition 3(ii) shows that greater dispersion in parental transfers increases the optimal aid amount when the EIS is not too small. This theoretical result continues to hold in our quantitative model, even though our estimated EIS is relatively low.

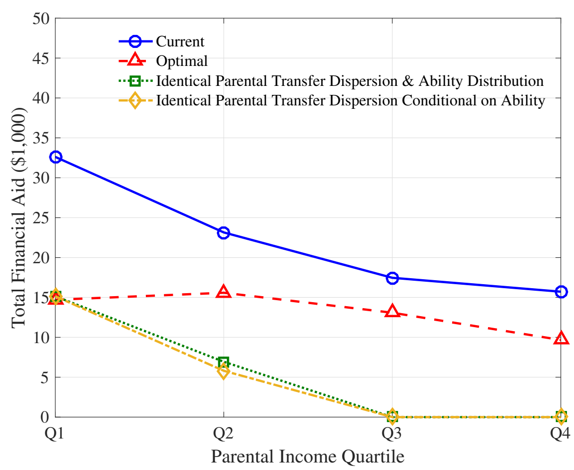
Together, these counterfactuals confirm that our theoretical insights on how the distribution of parental transfers shapes the progressivity of the optimal aid schedule carry over to the quantitative model: higher average parental transfers among higher-income families alone imply a steeply progressive aid schedule, but this effect is moderated by the greater dispersion in transfers for these families.

5.3 Expanding to the Full Schooling Choice Set

I now relax the binary choice restriction between attending a four-year college and not attending college at all, and instead allow for a full range of schooling choices, including two-year colleges. Parental transfers remain lump-sum and exogenous.



(a) 2-Year College



(b) 4-Year College

Figure 6: Financial Aid Schedules: Exogenous Parental Transfers

Figures 6(a) and 6(b) display the current (solid line) and budget-neutral optimal (dashed line) financial aid schedules for two-year and four-year colleges, respectively, with both the current and optimal policies spending approximately $7,000 per individual. Unlike in the binary choice case, the optimal policy departs significantly from the current one. Aid for four-year colleges is substantially reduced—especially for students in the bottom parental income quartile—while aid for two-year colleges increases, particularly for those from higher-income families. These changes make the system less progressive and shift public spending from four-year to two-year colleges.

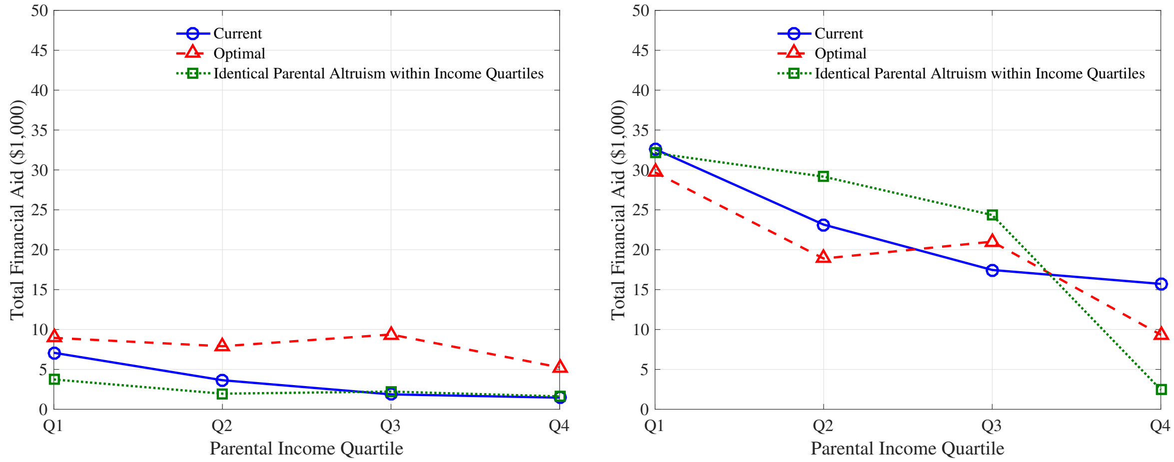
To understand the forces shaping the optimal aid schedule, I recalculate it under the two counterfactual distributions introduced earlier: one that varies only in average parental transfers across income quartiles, and another that also incorporates interquartile differences in ability. As in the previous exercise, the Lagrange multiplier on the constraint (21) is held fixed at its value under the actual distribution.

The dotted and dash-dotted lines in Figure 6 represent the optimal aid schedules under the first and second counterfactual distributions, respectively. Similar to the results in Figure 5 and consistent with Proposition 4(i), both counterfactual optimal policies are more progressive than the actual optimal policy, since the counterfactual distributions account only for higher average parental transfers, not the greater dispersion of transfers among higher-income parental quartiles. The optimal aid schedule under the second counterfactual (dash-dotted line) is very similar to the first, despite the positive correlation between parental income and ability. However, it is still substantially more progressive than the actual optimal policy (solid line), highlighting the role of varying parental transfer dispersion across income quartiles. As Proposition 4(ii) suggests, greater dispersion of parental transfers in higher-income groups increases optimal aid for these groups. This effect is especially pronounced for two-year colleges, for which the optimal aid schedule is effectively flat. While the optimal aid amount for four-year colleges exceeds that for two-year colleges among lower-income quartiles, the opposite is true for the top income quartile—an outcome consistent with Proposition 5(ii) and standing in stark contrast to the current policy, which allocates more aid to four-year colleges.

These results underscore that the impact of the parental transfer distribution on the optimal progressivity of financial aid depends on the level of schooling. Youth with lower parental transfers disproportionately self-select into more affordable two-year colleges, raising the marginal benefit of expanding aid at these institutions. This effect is particularly strong for higher-income groups with greater inequality in parental transfers, producing a relatively flat and elevated aid schedule for two-year colleges.

5.4 Endogenous Parental Transfers

Up to this point, I have focused on a simplified framework aligned with the theoretical model, where parents provide a lump-sum transfer under the current policy, and this amount is held fixed when evaluating alternative policies. I now turn to the full quantitative model, in which parental transfers vary across schooling levels and are endogenously determined based on the policy in place—not just the current one. In this setting, a \$1 increase in financial aid results in less than a \$1 increase in youth resources, as parents partially offset the aid by reducing their



(a) 2-Year College (b) 4-Year College

Figure 7: Financial Aid Schedules: Endogenous Parental Transfers

Table 2: Effects of Optimal vs. Current Policy under Endogenous Parental Transfers

	Parental Income Quartile				All
	Q1	Q2	Q3	Q4	
A. 2-Year Colleges Only					
Financial Aid ($1,000)	1.9	4.3	7.5	3.8	4.4
Parental Transfer ($1,000)	0.5	1.8	3.0	8.8	2.2
Attendance Rate (ppt)	1.4	4.4	6.3	5.5	4.4
Consumption Distortion ($1,000)	-2.7	-6.7	-12.6	-7.9	-7.5
B. 2- and 4-Year Colleges					
Financial Aid ($1,000)	0.5	2.2	9.2	0.7	3.4
Parental Transfer ($1,000)	-1.2	-6.2	-25	-19.2	-10.5
Attendance Rate (ppt)	1.0	1.0	1.5	0.9	1.1
Consumption Distortion ($1,000)	1.3	0.6	4.7	1.1	2.1
C. 4-Year Colleges Only					
Financial Aid ($1,000)	-2.9	-4.3	3.6	-6.4	-2.4
Parental Transfer ($1,000)	-3.7	-4.8	-15.7	1.3	-7.4
Attendance Rate (ppt)	-1.8	-3.3	-2.0	-5.5	-3.1
Consumption Distortion ($1,000)	0.5	2.9	-0.2	1.6	1.1
D. All Individuals					
Financial Aid ($1,000)	-0.4	-0.5	3.6	-2.8	0.0
Parental Transfer ($1,000)	-0.2	-0.9	-2.9	0.0	-1.0
College Attendance Rate (ppt)	0.6	2.1	5.8	0.9	2.3
Efficiency Loss ($1,000)	-0.2	-0.7	-2.5	-0.6	-1.0
Consumption Distortion	-0.1	0.3	-0.4	0.4	0.0
Schooling Distortion	-0.1	-1.0	-2.1	-1.0	-1.0
Welfare ($1,000)	-0.4	-0.7	3.2	-2.1	0.0

transfers. That is, the marginal utility of additional aid is dampened by the crowding out effect. Naturally, this effect is weaker for youth who receive little parental support to begin with.

The dashed lines in Figure 7 show the optimal aid schedules with endogenous parental transfers, which allocate approximately $7,900 per individual. Relative to the current policy (solid lines), which spends the same amount, the optimal policy for four-year colleges is only slightly more progressive in terms of the aid gap between the bottom and top parental income quartiles, with generally lower aid amounts—except for the third quartile. For two-year colleges, the optimal aid schedule is both higher and less progressive than the current one. Compared to the case with exogenous parental transfers (Figure 6), the optimal aid schedules are steeper when transfers are endogenous. This reflects the fact that the crowding out of parental transfers is stronger for youth in higher-income quartiles, who receive larger transfers on average.

While endogenous parental transfers make the optimal financial aid schedules more progressive, the greater dispersion of transfers among youth from higher-income families can still contribute to a flattening of the aid schedules. To evaluate this effect, I consider a counterfactual environment that abstracts from within-group heterogeneity in parental altruism—the key driver of within-group variation in transfers. The budget-neutral optimal aid schedules under this counterfactual are shown as the dotted lines in Figure 7. For four-year colleges, the aid schedules are steeper, consistent with the reduced variation in within-group inequality across income quartiles when parental altruism is homogeneous. Although progressivity is slightly lower for two-year colleges under the counterfactual, the overall level of aid is also lower, reflecting a reallocation of resources from two-year to four-year colleges. These results highlight the importance of accounting for within-group heterogeneity in parental transfers—driven by variation in altruism—when designing optimal financial aid policies.

Table 2 presents the effects of the optimal policy relative to the current policy in the baseline environment featuring within-group heterogeneity in parental altruism. Panels A–C report outcomes by schooling choice, while Panel D summarizes results for all individuals. As discussed earlier, the optimal policy increases financial aid for two-year colleges and reduces aid for four-year colleges, though to a lesser extent. As a result, individuals who attend both two- and four-year colleges receive more aid overall. These adjustments to the aid schedule lead to an increase in the share of individuals choosing two-year colleges, accompanied by a decline in those selecting the four-year-only option. Overall, the college attendance rate rises across all income groups under the optimal policy.

The optimal policy reduces overall efficiency losses by $1,000 per individual, entirely driven by a decline in the schooling distortion. In contrast, average consumption distortion remains largely unchanged, despite a more efficient allocation of aid. The substantial increase in financial aid for two-year colleges across all parental income quartiles significantly reduces consumption distortion among individuals who attend only two-year colleges. However, their consumption allocation remains distorted on average, so the increase in their attendance rate under the optimal policy places upward pressure on the average consumption distortion, partially offsetting improvements elsewhere.

Efficiency losses decline for all parental income quartiles, despite most groups receiving less financial aid and parental transfers on average under the optimal policy—only the third parental income quartile experiences an increase in aid sufficient to offset the reduction in parental transfers. The reduction in schooling distortion appears to be the primary driver of efficiency gains across all groups, while the intertemporal consumption allocation worsens for the second and fourth quartiles—those experiencing the largest reductions in financial aid.

Youth's welfare gains from the reform, measured by the change in their money-metric indirect utility, are equal to the sum of changes in financial aid and parental transfers, plus the efficiency gain (Lemma 3 for the case of endogenous parental transfers). Under the optimal policy, aggregate welfare does not improve, as the efficiency gain of \$1,000 per individual is fully offset by the crowding out of parental transfers.[35] At the group level, welfare gains largely mirror changes in financial aid, with gains accruing only to the third parental income quartile.

These results indicate that simply redesigning the financial aid schedules within the existing budget could generate meaningful efficiency gains by better targeting individuals more likely to be borrowing constrained due to low parental transfers. However, such efficiency gains do not necessarily lead to improvements in youth welfare if parents respond by reducing their transfers.

6 Externality of Education

Until now, the analysis has focused on financial aid as a tool to address inefficiencies caused by borrowing constraints. However, a complementary rationale for subsidizing education is the presence of externalities—that the social return to education exceeds the private return. In section, I extend the analysis to the case where private returns are suppressed by the disincentive effects of income taxation on education (Bovenberg and Jacobs, 2005).

6.1 Modifying the Optimal Policy Formula

I first show how the optimal policy formula (6) is modified to account for externalities in the theoretical framework of Section 3.3, which features binary choice and exogenous parental transfers. For $j \in \{0, 1\}$, let y_j, defined earlier, represent after-tax earnings, and define $\tilde{y}_j$ as before-tax earnings, so that $\tau_j := \tilde{y}_j - y_j$ is the income tax paid. I assume $\tau_1 > \tau_0$, implying that the private earnings returns to college, $y_1 - y_0$, is smaller than the social return, $\tilde{y}_1 - \tilde{y}_0$.

[35] When family welfare is defined as the sum of (13) and (14), it increases by \$1,000 due to higher parental consumption resulting from reduced transfers. With endogenous transfers, minimizing inefficiencies for a given level of aid spending is equal to maximizing family welfare. See Section D.1 of the Appendix for details.

While individuals base their schooling decisions on after-tax earnings, the social planner values before-tax earnings. Let $\tilde{L}_s(b,g)$ denote the efficiency loss from suboptimal schooling choices, accounting for the social return. It mirrors $L_s(b,g)$, but replaces y_j with $\tilde{y}_j$. The total efficiency loss, $\tilde{L}(b,g) := \tilde{L}_s(b,g) + L_c(b,g)$ can then be written, up to constant terms, as $L(b,g)$ minus the total tax paid under the current allocation, $[1 - p(b,g)]\tau_0 + p(b,g)\tau_1$.[36]

The social planner allocates income tax revenues and other funds to financial aid and other expenditures. The planner's budget constraint is given by:

$$\underbrace{\sum_{h\in\mathcal{H}} n_h P_h(g_h) g_h}_{\text{spending on financial aid}} + \underbrace{E}_{\text{other net spending}} \leq \underbrace{\sum_{h\in\mathcal{H}} n_h \Big\{\big[1 - P_h(g_h)\big]\tau_0 + P_h(g_h)\tau_1\Big\}}_{\text{revenue from income tax}}, \tag{22}$$

where $E > 0$ represents net non-aid spending and is assumed to be exogenous. Unlike the previous budget constraint (4), the planner's spending on financial aid is not fixed: additional aid can be funded through higher tax revenues.

Taking tax policies as given, the optimal financial aid schedule that accounts for the education externality, denoted $(\tilde{g}_h)_{h\in\mathcal{H}}$, minimizes the total efficiency loss $\tilde{L}(b,g)$, aggregated across individuals, subject to (22) and $g_h \geq 0$ for all h. When $\tilde{g}_h > 0$, the optimal aid satisfies

$$\tilde{g}_h = \underbrace{\tau_1 - \tau_0}_{\text{Pigouvian correction}} + \frac{P_h(\tilde{g}_h)}{P'_h(\tilde{g}_h)} \left\{ \frac{\int V'_1(b + \tilde{g}_h) m_{h|1}(b, \tilde{g}_h) db}{1 + \tilde{\lambda}} - 1 \right\}, \tag{23}$$

where $\tilde{\lambda} \geq 0$ is the Lagrange multiplier on the constraint (22). This expression is identical to the baseline formula (6), except for the Pigouvian correction term $\tau_1 - \tau_0$, which increases the optimal level of aid to offset the effects of taxation, as shown by Bovenberg and Jacobs (2005).

When $\tau_1 = \tau_0$, there is no incentive to encourage college attendance at the margin, due to a *negative* fiscal externality: aid given to marginal students—those indifferent between attending and not attending—raises costs without yielding benefits. In contrast, when $\tau_1 > \tau_0$, expanding college attendance can be socially beneficial if it increases tax revenues net of aid costs. In this case, total aid spending can also rise, provided it is offset by higher tax revenues.

Although the Pigouvian correction term is identical across groups, it can give rise to a progressive aid schedule due to heterogeneity in behavioral responses, as highlighted by Colas, Findeisen, and Sachs (2021). To isolate this mechanism, consider the case where $\tilde{\lambda} \to \infty$, so that the planner maximizes net tax revenue. In this case, the optimal aid policy simplifies to $\tilde{g}_h = \tau_1 - \tau_0 - P_h(\tilde{g}_h)/P'_h(\tilde{g}_h)$, implying that groups with stronger enrollment responses—typically those with lower parental income—receive greater financial aid. This contrasts with the case without externalities, discussed in Section 3.3, where stronger enrollment responses increase the marginal cost of aid, thereby reducing optimal aid levels.

[36]See Section E.1 of the Appendix for the expressions of $\tilde{L}_s(b,g)$ and $\tilde{L}(b,g)$.

Nonetheless, the optimal policy formula (23) shows that the paper's central mechanism remains relevant with externality: variation in the distribution of parental transfers across groups shapes the optimal aid schedule by altering the marginal benefit of aid.

6.2 Quantifying the Effects of Externality

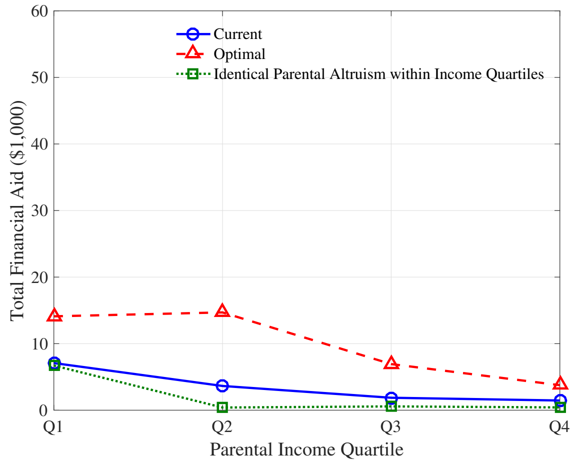
I now solve for the optimal policy accounting for externality in the quantitative model from Section 4.[37] The dashed lines in Figure 8 show the optimal financial aid schedules when accounting for the externality, and Table 3 reports the effects of the optimal policy relative to the current policy. Under the optimal policy, total spending on financial aid increases by approximately $5,000 per individual, fully financed by higher tax revenues from greater educational attainment. This expansion is concentrated among the bottom two income quartiles, making the aid schedule more progressive than the current policy for both two-year and four-year colleges. The progressivity is particularly pronounced for four-year colleges: students in the top parental income quartile receive less aid than under the current policy, despite the overall increase in program size.[38] Aid also expands more at two-year colleges than at four-year colleges, with two-year college students receiving more support than under the current policy across all income groups.

The optimal aid schedules are also more progressive than in the case without externality (Figure 7). As discussed in Section 6.1, this reflects that college enrollment among students from lower parental income groups tends to respond more strongly to aid expansion, generating greater tax revenues. This strong progressivity induced by the fiscal externality can be moderated by greater inequality in parental transfers among higher-income groups. The dotted lines in Figure 8 show the optimal financial aid schedules under the same externality, but in the counterfactual environment considered earlier, where parental altruism is homogeneous within each income quartile. These schedules are more progressive for both college types than those under the baseline environment (dashed lines). Moreover, optimal aid amounts for two-year colleges in the counterfactual are lower than those in the baseline and even lower than under the current policy. These patterns mirror the results without externality, confirming a key insight from the quantitative exercise: within-group heterogeneity in parental altruism—which drives greater inequality in parental transfers among higher-income groups—reduces the progressivity of optimal aid and reallocates resources toward two-year colleges.

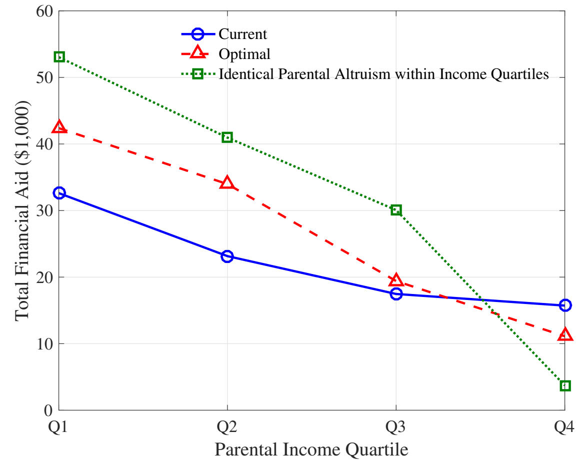
Returning to the baseline environment, the optimal financial aid schedules yield a substantial reduction in inefficiencies—approximately $7,000 per individual. Across all parental income quartiles, this improvement is driven entirely by a better allocation of schooling investment,

[37] See Section E.2 of the Appendix for the modified definitions of schooling distortion and the planner's problem that incorporate externality.

[38] The progressivity of the optimal financial aid schedule for four-year colleges is quantitatively comparable to the baseline result of Colas, Findeisen, and Sachs (2021).



(a) 2-Year College



(b) 4-Year College

Figure 8: Financial Aid Schedules: Endogenous Parental Transfers and Externality

Table 3: Effects of Optimal vs. Current Policy under Endogenous Parental Transfers and Externality

	Parental Income Quartile				All
	Q1	Q2	Q3	Q4	
A. 2-Year Colleges Only					
Financial Aid ($1,000)	7.0	11.0	5.1	2.3	6.2
Parental Transfer ($1,000)	-3.9	-14.1	-19.5	-10.6	-11.4
Attendance Rate (ppt)	0.9	4.5	4.7	3.2	3.3
Consumption Distortion ($1,000)	-9.9	-15.7	-9.0	-4.9	-9.6
B. 2- and 4-Year Colleges					
Financial Aid ($1,000)	11.8	16.3	6.0	0.1	9.8
Parental Transfer ($1,000)	-17.0	-30.5	-8.9	0.5	-28.6
Attendance Rate (ppt)	3.4	3.7	1.0	0.6	2.2
Consumption Distortion ($1,000)	-3.8	-0.1	2.7	-0.3	1.0
C. 4-Year Colleges Only					
Financial Aid ($1,000)	9.8	10.8	1.9	-4.6	5.0
Parental Transfer ($1,000)	-40.3	-36.7	-1.3	4.1	-29.5
Attendance Rate (ppt)	14.4	9.7	-0.9	-2.8	5.1
Consumption Distortion ($1,000)	4.2	4.4	0.6	1.4	3.3
D. All Individuals					
Financial Aid ($1,000)	10.4	10.4	2.2	-2.1	5.2
Income Tax ($1,000)	9.8	9.1	2.0	0.0	5.2
Parental Transfer ($1,000)	0.4	-1.2	-1.6	1.0	-0.4
College Attendance Rate (ppt)	18.6	17.9	4.8	1.1	10.6
Efficiency Loss ($1,000)	-11.1	-12.8	-3.9	-0.7	-7.1
Consumption Distortion	3.1	1.5	0.2	0.5	1.3
Schooling Distortion	-14.2	-14.3	-4.1	-1.2	-8.4
Welfare ($1,000)	12.2	12.9	2.5	-0.4	6.8

while intertemporal consumption allocation exacerbates, due to the significant increase in educational attainment induced by the aid expansion. Notably, the increased aid amounts for four-year colleges worsen intertemporal consumption smoothing among those who attend only four-year colleges. This deterioration results from a reduction in parental transfers, which reflects both the crowding out effect and a compositional shift, the latter arising as expanded aid at four-year colleges draws in individuals with lower parental transfers into those institutions.

With the externality, the youth's welfare gain from the optimal policy equals the sum of (i) changes in financial aid, *net of taxes paid*—the only difference from the case without externality—(ii) changes in parental transfers, and (iii) the efficiency gain.[39] The optimal policy improves youth welfare by around $7,000 per individual—an amount roughly equal to the efficiency gain, as the crowd out of parental transfers is relatively small. These gains primarily accrue to the bottom two quartiles, where financial aid increases the most. Because the increase in aid for each group is largely offset by higher taxes paid by that group, and parental transfer adjustments are relatively small, the welfare improvements across income quartiles primarily reflect efficiency gains, particularly from better schooling allocation.

In sum, recognizing the fiscal externality of education strengthens the case for progressive and expansive financial aid policies. Even when accounting for behavioral responses by parents, well-designed aid expansions can yield substantial efficiency and welfare gains. Importantly, the central insight from the baseline analysis without externality remains: greater inequality in parental transfers among higher-income groups, driven by heterogeneity in parental altruism, weakens the progressivity of optimal aid and shifts resources toward two-year colleges. In fact, absent this heterogeneity, the optimal aid schedules would be even more progressive.

7 Conclusions

As the cost of higher education continues to rise, there is increasing concern that many young people without substantial parental support lack the necessary resources to attend college. While current policy discussions around financial aid mainly focus on disparities in family resources, this paper highlights significant inequality in parental transfers even among students from families with similar financial means. This poses a challenge for financial aid design, as parental transfers are not directly observable by aid authorities, complicating efforts to target assistance effectively.

I show that unobservable heterogeneity in parental transfers—conditional on family resources—has important implications for the design of need-based financial aid, even when equity concerns are set aside. Youth who receive low parental transfers are more likely to be borrowing constrained and to underinvest in education, and can therefore be better targeted by increasing aid at lower levels of schooling, such as two-year colleges. Because parental transfers

[39] See Section E.2 in the Appendix for details.

are only imperfectly correlated with family resources, financial aid need not decline sharply with family resources. The quantitative analysis suggests a reform that reallocates financial aid from four-year to two-year colleges to improve targeting. Such a reform also reduces the gap in aid across income groups for two-year colleges, thereby offering broader public support for college attendance—including for youth from high-income families. The potential efficiency gains from such a budget-neutral reform, while modest, are nontrivial.

Universally expanding public support for two-year colleges has been a prominent topic in policy debates, particularly since President Obama introduced America's College Promise in 2015—a legislative initiative to make community college tuition-free. While free community college proposals have drawn substantial media attention and led to legislation at the state and local levels, critics argue that such policies are poorly targeted: "covering the full tuition of all community college students would mean middle-income, and even upper-income, students would get hefty subsidies, even though many do not need the help" (Butler, 2015).

However, the results in this paper suggest that eliminating community college tuition could improve efficiency, as some students from middle- and high-income families receive little parental support and would benefit from additional aid. At the same time, the analysis implies that generous aid for two-year colleges should be offset by reduced aid for four-year colleges, since students with substantial parental support—who are less in need of aid—are disproportionately likely to attend four-year institutions. Therefore, free community college proposals should be considered within a broader framework that reevaluates tuition and financial aid policies across all types of institutions.

References

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L. Violante. 2019. "Education Policy and Intergenerational Transfers in Equilibrium." *Journal of Political Economy* 127 (6):2569–2624.

Becker, Gary S. 1975. *Human Capital*. New York, NY: Columbia University Press.

Becker, Gary S. and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4 (3):S1–S39.

Belley, Philippe and Lance Lochner. 2007. "The Changing Role of Family Income and Ability in Determining Educational Achievement." *Journal of Human Capital* 1 (1):37–89.

Bénabou, Roland. 2002. "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica* 70 (2):481–517.

Bernheim, B. Douglas, Andrei Shleifer, and Lawrence H. Summers. 1985. "The Strategic Bequest Motive." *Journal of Political Economy* 93 (6):1045–1076.

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

Bovenberg, A. Lans and Bas Jacobs. 2005. "Redistribution and Education Subsidies are Siamese Twins." *Journal of Public Economics* 89 (11):2005–2035.

Brown, Meta, John Karl Scholz, and Ananth Seshadri. 2012. "A New Test of Borrowing Constraints for Education." *Review of Economic Studies* 79 (2):511–538.

Browning, Martin, Lars Peter Hansen, and James J. Heckman. 1999. "Micro Data and General Equilibrium Models." In *Handbook of Macroeconomics*, vol. 1, Part A, edited by John B. Taylor and Michael Woodford, chap. 8. Elsevier, 543–633.

Bruce, Neil and Michael Waldman. 1990. "The Rotten-Kid Theorem Meets the Samaritan's Dilemma." *Quarterly Journal of Economics* 105 (1):155–165.

Butler, Stuart M. 2015. "Obama's SOTU Free College Plan is Bad for Poor Americans." https://www.brookings.edu/opinions/why-obamas-free-community-college-idea-is-a-bad-one/. (accessed: July 1, 2025).

Caucutt, Elizabeth M., Lance Lochner, and Youngmin Park. 2017. "Correlation, Consumption, Confusion, or Constraints: Why Do Poor Children Perform So Poorly?" *Scandinavian Journal of Economics* 119 (1):102–147.

Colas, Mark, Sebastian Findeisen, and Dominik Sachs. 2021. "Optimal Need-Based Financial Aid." *Journal of Political Economy* 129 (2):492–533.

College Board. 2016. *Trends in Student Aid 2016*. New York, NY: College Board.

Consumer Financial Protection Bureau. 2012. *Private Student Loans*. Washington, DC: Consumer Financial Protection Bureau.

Córdoba, Juan Carlos and Marla Ripoll. 2018. "The Elasticity of Intergenerational Substitution, Parental Altruism, and Fertility Choice." *The Review of Economic Studies* 86 (5):1935–1972.

Costinot, Arnaud and Jonathan Vogel. 2010. "Matching and Inequality in the World Economy." *Journal of Political Economy* 118 (4):pp. 747–786.

De Fraja, Gianni. 2002. "The Design of Optimal Education Policies." *Review of Economic Studies* 69 (2):437–466.

Dick, Andrew W. and Aaron S. Edlin. 1997. "The Implicit Taxes from College Financial Aid." *Journal of Public Economics* 65 (3):295–322.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64 (5):1001–1044.

Do, Chau. 2004. "The Effects of Local Colleges on the Quality of College Attended." *Economics of Education Review* 23 (3):249–257.

Ellwood, David T. and Thomas J. Kane. 2000. "Who Is Getting a College Education? Family Background and the Growing Gaps in Enrollment." In *Securing the Future: Investing in Children from Birth to College*, edited by Sheldon Danziger and Jane Waldfogel. New York: Russell Sage Foundation, 283–324.

Findeisen, Sebastian and Dominik Sachs. 2016. "Education and Optimal Dynamic Taxation:

The Role of Income-Contingent Student Loans." *Journal of Public Economics* 138:1–21.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura. 2014. "Income Taxation of U.S. Households: Facts and Parametric Estimates." *Review of Economic Dynamics* 17 (4):559–581.

Haider, Steven J. and Kathleen M. McGarry. 2012. "Parental Investments in College and Later Cash Transfers." Working Paper 18485, National Bureau of Economic Research.

Hanushek, Eric A., Charles Ka Yui Leung, and Kuzey Yilmaz. 2003. "Redistribution through Education and Other Transfer Mechanisms." *Journal of Monetary Economics* 50 (8):1719–1750.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." In *Handbook of the Economics of Education*, vol. 1, edited by E. Hanushek and F. Welch. Elsevier, 307–458.

Hopkins, Ed and Tatiana Kornienko. 2004. "Running to Keep in the Same Place: Consumer Choice as a Game of Status." *American Economic Review* 94 (4):1085–1107.

Hu, Yingyao. 2008. "Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution." *Journal of Econometrics* 144 (1):27–61.

Hu, Yingyao and Susanne M. Schennach. 2008. "Instrumental Variable Treatment of Nonclassical Measurement Error Models." *Econometrica* 76 (1):195–216.

Jacobs, Bas and Hongyan Yang. 2016. "Second-best income taxation and education policy with endogenous human capital and borrowing constraints." *International Tax and Public Finance* 23:234–268.

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

Kaplan, Greg. 2012. "Moving Back Home: Insurance against Labor Market Risk." *Journal of Political Economy* 120 (3):446–512.

Keane, Michael P. and Kenneth I. Wolpin. 2001. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review* 42 (4):1051–1103.

Kehoe, Timothy J. and David K. Levine. 1993. "Debt-Constrained Asset Markets." *Review of Economic Studies* 60 (4):865–888.

Krueger, Dirk and Alexander Ludwig. 2016. "On the Optimal Provision of Social Insurance: Progressive Taxation versus Education Subsidies in General Equilibrium." *Journal of Monetary Economics* 77:72–98.

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

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

McKenzie, Lionel. 1957. "Demand Theory Without a Utility Index." *Review of Economic*

Studies 24 (3):185–189.

McPherson, Michael S. and Morton Owen Schapiro. 1998. *The Student Aid Game: Meeting Need and Rewarding Talent in American Higher Education.* Princeton University Press.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics* 12 (2):380–391.

Ramos, Hector M, Jorge Ollero, and Miguel A Sordo. 2000. "A Sufficient Condition for Generalized Lorenz Order." *Journal of Economic Theory* 90 (2):286–292.

Rothschild, Michael and Joseph E. Stiglitz. 1970. "Increasing risk: I. A definition." *Journal of Economic Theory* 2 (3):225–243.

———. 1971. "Increasing risk II: Its economic consequences." *Journal of Economic Theory* 3 (1):66–84.

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

Sallie Mae. 2024. *How America Pays for College 2024*. Newark, DE: Sallie Mae Bank.

Stantcheva, Stefanie. 2017. "Optimal Taxation and Human Capital Policies over the Life Cycle." *Journal of Political Economy* 125 (6):1931–1990.

Appendices

A Additional Details on the NPSAS:12 Data

Parental transfer amounts are based on the student survey question: "Through the end of the 2011–12 school year, about how much will your parents (or guardians) have helped you pay for any of your education and living expenses while you are enrolled in school?" The responses are reported in 12 categories,[1] from which I compute percentiles by assume a uniform distribution within each interval. To calculate the mean reported in Table A.1, I additionally assume a Pareto distribution at the top two bins and estimate the shape parameter based on the bin shares.

Table A.1: Average College Costs and Financing Sources ($)

Variable	2-Year	4-Year
Tuition and Fees	2,846	8,198
Cost of Attendance	13,974	22,336
Expected Family Contribution	6,769	12,552
Parental Transfers	2,065	5,884
Earnings While Enrolled	4,164	2,673
Grants	3,911	5,771
Federal Student Loans	1,273	4,647
Private Student Loans	110	524

Table A.1 presents average college costs and financing sources for the sample.[2] In the 2011–2012 academic year, the average cost of attendance was $13,974 for two-year colleges and $22,336 for four-year colleges, with the difference driven largely by tuition and fees. On average, about half of the cost is expected to be covered by students and their families through the EFC, while grants account for roughly 30%, and the remaining gap can be financed through federal student loans.[3] In practice, however, families often contribute less than their EFC, with parents covering less than half of the EFC on average. Even after including students' own earnings, this falls short of meeting the full cost of attendance—particularly at four-year institutions. The gap is filled in part through federal student loans and, presumably, by reducing living expenses.[4]

[1]The categories are: $0, less than $250, $250 to $500, $501 to $1,000, $1,001 to $1,500, $1,501 to $2,000, $2,001 to $5,000, $5,001 to $10,000, $10,001 to $15,000, $15,001 to $20,000, $20,001 to $25,000, and more than $25,000.

[2]All results based on the NPSAS:12 data use sample weights and are obtained through the National Center for Education Statistics DataLab website. Direct access to the data is granted only to users in the U.S.

[3]For dependent students in 2011–2012, the annual loan limits were $5,500 for first-year students, $6,500 for second-year students, and $7,500 for third-year students and beyond.

[4]Students rarely borrowed from private lenders, likely due to higher costs and more stringent requirements.

B Proofs and Analytical Details

B.1 Money-Metric Utility

I show that the utility function (1) is a money-metric transformation of the standard additively separable utility function $u(c_1) + \beta u(c_2)$, where $u(c) = (c^{1-1/\gamma} - 1)/(1 - 1/\gamma)$. Define the expenditure function that gives the minimum lifetime consumption to achieve a given utility level $\tilde{U}$:

$$\begin{aligned} \underline{C}(\tilde{U}, R) &:= \min_{c_1, c_2} \left\{ c_1 + R^{-1} c_2 \right\} \\ &\text{subject to } u(c_1) + \beta u(c_2) \geq \tilde{U}, \\ &c_1 \geq 0, c_2 \geq 0. \end{aligned}$$

Then, $\underline{C}\big(u(c_1) + \beta u(c_2), R\big)$ is the money-metric utility function that measures individual welfare in monetary terms (McKenzie, 1957). Since $\beta R = 1$, it is equal to the utility function (1):

$$\begin{aligned} \underline{C}\big(u(c_1) + \beta u(c_2), R\big) = &\left\{ \frac{\left(1 - \frac{1}{\gamma}\right) \left[u(c_1) + \beta u(c_2)\right] + (1 + \beta)}{\left[1 + \left(\beta R^{1-\frac{1}{\gamma}}\right)^{\gamma}\right]^{\frac{1}{\gamma}}} \right\}^{\frac{\gamma}{\gamma-1}} \\ =&(1+\beta) \left(\frac{1}{1+\beta} c_1^{\frac{\gamma-1}{\gamma}} + \frac{\beta}{1+\beta} c_2^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}} \\ =&U(c_1, c_2). \end{aligned}$$

B.2 Proof of Lemma 1

When (2) does not bind, $V_1(b+g) = y_1 + b + g - k$ and $V_1'(b+g) = 1$. When (2) binds, $V_1(b+g) = U(c_1, c_2)$, where $c_1 = b + g - k + \bar{l} < c_2 = R(y_1 - \bar{l})$. Therefore,

$$V_1'(b+g) = \left(\frac{1}{1+\beta} + \frac{\beta}{1+\beta} \left(\frac{c_2}{c_1} \right)^{\frac{\gamma-1}{\gamma}} \right)^{\frac{1}{\gamma-1}} > 1.$$

B.3 Proof of Lemma 2

Part (i): The derivative of $p(b, g)$ with respect to b is $f\big(V_1(b+g) - V_0(b)\big)\big[V_1'(b+g) - V_0'(b)\big]$. Therefore, the result follows from $V_0'(b) = 1$ and Lemma 1.

Part (ii): The derivative of $p(b, g)$ with respect to g is $f\big(V_1(b+g) - V_0(b)\big)V_1'(b+g) \geq 0$.

B.4 Proof of Lemma 3

First, notice that $L_s(b,g)$ can be written as follows:

$$L_s(b,g) = \int \max\{y_0+\varepsilon, y_1-k\}dF(\varepsilon) - V(b,g) - \big[1-p(b,g)\big]\big[y_0 - V_0(b)\big] \\ - p(b,g)\big[y_1 - k - V_1(b+g)\big],$$

which follows from

$$\begin{aligned} &\mathbb{I}_{\varepsilon > V_1(b+g)-V_0(b)}(y_0+\varepsilon) + \mathbb{I}_{\varepsilon \le V_1(b+g)-V_0(b)}(y_1-k) \\ =&\mathbb{I}_{\varepsilon > V_1(b+g)-V_0(b)}\big[V_0(b)+\varepsilon+y_0-V_0(b)\big]\mathbb{I}_{\varepsilon \le V_1(b+g)-V_0(b)}\big[V_1(b+g)+y_1-k-V_1(b+g)\big] \\ =&\mathbb{I}_{\varepsilon > V_1(b+g)-V_0(b)}\big[V_0(b)+\varepsilon\big] + \mathbb{I}_{\varepsilon \le V_1(b+g)-V_0(b)}V_1(b+g) \\ &+ \mathbb{I}_{\varepsilon > V_1(b+g)-V_0(b)}\big[y_0 - V_0(b)\big] + \mathbb{I}_{\varepsilon \le V_1(b+g)-V_0(b)}\big[y_1-k-V_1(b+g)\big]. \end{aligned}$$

Therefore,

$$\begin{aligned} L_s(b,g) + L_c(b,g) =& \int \max\{y_0+\varepsilon, y_1-k\}dF(\varepsilon) - V(b,g) + \big[1-p(b,g)\big]b \\ &- p(b,g)\big[y_1-k-V_1(b+g)\big] + p(b,g)\big[(y_1+b+g-k) - V_1(b+g)\big] \\ =& \int \max\{y_0+\varepsilon, y_1-k\}dF(\varepsilon)b + p(b,g)g - V(b,g). \end{aligned}$$

B.5 Proof of Proposition 1

Suppose that (2) does not bind for all individuals in group h and $\hat{g}_h > 0$. Then the first order condition becomes

$$1 = (1+\lambda)\left[1 + \frac{P_h'(\hat{g}_h)}{P_h(\hat{g}_h)}\hat{g}_h\right].$$

This condition can hold only when $P_h'(\hat{g}_h) = \lambda = 0$. If $P_h'(\hat{g}_h) > 0$, then the right hand side of the equation is strictly greater than 1. Therefore, $\hat{g}_h = 0$ must hold in this case.

B.6 Proof of Corollary 1

Suppose that (2) does not bind for all individuals. If $\hat{g}_h = 0$ for all h, then $\sum_h n_h P_h(\hat{g}_h)\hat{g}_h = 0 < G$, implying $\lambda = 0$. If $\hat{g}_h > 0$ for some h, then $\lambda = 0$ must hold by Proposition 1.

B.7 Proof of Lemma 4

Since $V_1'(b+g) = 1$ when (2) is slack, we only need to consider the case (2) is binding. In this case, the derivatives of the indirect utility function are partial derivatives of the utility function with respect to c_1.

The first derivative is

$$\frac{\partial U(c_1, c_2)}{\partial c_1} = \left[\frac{1}{1+\beta} + \frac{\beta}{1+\beta} c_1^{-\frac{\gamma-1}{\gamma}} c_2^{\frac{\gamma-1}{\gamma}} \right]^{\frac{1}{\gamma-1}}.$$

The second derivative is

$$\frac{\partial^2 U(c_1, c_2)}{\partial c_1^2} = -\frac{1}{\gamma} \frac{\beta}{1+\beta} \left[\frac{1}{1+\beta} c_1^{\frac{\gamma-1}{\gamma} \frac{1-2\gamma}{2-\gamma}} c_2^{\frac{(\gamma-1)^2}{\gamma(2-\gamma)}} + \frac{\beta}{1+\beta} c_1^{\frac{\gamma-1}{\gamma}\left(\frac{1+\gamma}{\gamma-2}\right)} c_2^{\frac{\gamma-1}{\gamma(2-\gamma)}} \right]^{\frac{2-\gamma}{\gamma-1}} < 0.$$

The third derivative is

$$\frac{\partial^3 U(c_1, c_2)}{\partial c_1^3} = \frac{1}{\gamma^2} \frac{\beta}{(1+\beta)^2} \left[\frac{1}{1+\beta} c_1^{\frac{\gamma-1}{\gamma} \frac{1-2\gamma}{2-\gamma}} c_2^{\frac{(\gamma-1)^2}{\gamma(2-\gamma)}} + \frac{\beta}{1+\beta} c_1^{\frac{\gamma-1}{\gamma}\left(\frac{1+\gamma}{\gamma-2}\right)} c_2^{\frac{\gamma-1}{\gamma(2-\gamma)}} \right]^{\frac{3-2\gamma}{\gamma-1}} \times$$
$$\left[(2\gamma-1) c_1^{\frac{-\gamma^2+\gamma-1}{\gamma(2-\gamma)}} c_2^{\frac{(\gamma-1)^2}{\gamma(2-\gamma)}} + (1+\gamma)\beta c_1^{\frac{-(2\gamma-1)}{\gamma(2-\gamma)}} c_2^{\frac{\gamma-1}{\gamma(2-\gamma)}} \right],$$

which is positive if $\gamma \geq 1/2$.

B.8 Proof of Lemma 5

First, notice that $m_{h'|1}(b,g)/m_{h|1}(b,g) = [p(b,g)m_{h'}(b)/P_{h'}(g)]/[p(b,g)m_h(b)/P_h(g)] = [m_{h'}(b)/P_{h'}(g)]/[m_h(b)/P_h(g)]$ holds. Therefore, $m_{h'|1}(b,g)/m_{h|1}(b,g)$ is also increasing in b. Since the monotone likelihood ratio property implies first-order stochastic dominance (Milgrom, 1981), the parental transfer distribution conditional on college attendance of group h' first-order stochastically dominates that of group h. Since $V_1'(b+g)$ is decreasing (Lemma 4), $\int V_1'(b+g)m_{h'|1}(b,g)db \leq \int V_1'(b+g)m_{h|1}(b,g)db$ holds for all g.

B.9 Proof of Lemma 6

Since $m_{h|1}(b,g)/m_{h'|1}(b,g) = [m_h(b)/P_h(g)]/[m_{h'}(b)/P_{h'}(g)]$, the unimodality of $m_h(b)/m_{h'}(b)$ implies that $m_{h|1}(b,g)/m_{h'|1}(b,g)$ is also unimodal in b. As shown by Ramos, Ollero, and Sordo (2000) and Hopkins and Kornienko (2004), the unimodality of $m_{h|1}(b,g)/m_{h'|1}(b,g)$, along with $\int b\left[m_{h|1}(b,g) - m_{h'|1}(b,g)\right]db = 0$, implies that the parental transfer distribution conditional on college attendance of group h second-order stochastically dominates that of group h'. Since $V_1'(b+g)$ is convex when $\gamma \geq 1/2$ (Lemma 4), $\int V_1'(b+g)m_{h'|1}(b,g)db \geq \int V_1'(b+g)m_{h|1}(b,g)db$ holds.

B.10 Proof of Proposition 2

First, note that $\mathrm{E}[z] = -1$ and $\mathrm{SD}(z) = 1$. Therefore, the support of ε is bounded above by $\mu + \sigma$. The cumulative distribution function of ε over its support is

$$F(\varepsilon) = \Pr\left(\mu + \sigma(z+1) \le \varepsilon\right) = \Pr\left(z \le \frac{\varepsilon - \mu}{\sigma} - 1\right) = \exp\left(\frac{\varepsilon - \mu}{\sigma} - 1\right),$$

and its density is

$$f(\varepsilon) = \frac{1}{\sigma}\exp\left(\frac{\varepsilon - \mu}{\sigma} - 1\right) = \frac{F(\varepsilon)}{\sigma}.$$

Using the expression for the semi-elasticity (7) for the case of $p(b, \hat{g}_h) < 1$ for all b, the optimal policy formula (6) becomes (8).

B.11 Proof of Lemma 7

Consider a fixed group h and define $x := b + g$. Let $\tilde{m}_{h|1}(x, g)$ denote the density of x conditional on g among college attendees. It is equal to $m_{h|1}(x - g, g)$ when $x \ge g$ and is zero otherwise.

I first show that, for $g' > g$, $\tilde{m}_{h|1}(x, g')/\tilde{m}_{h|1}(x, g) = m_{h|1}(x - g', g')/m_{h|1}(x - g, g)$ increases with x when $x \ge g'$. For $x' > x \ge g'$,

$$\frac{m_{h|1}(x' - g, g)}{m_{h|1}(x - g, g)} = \frac{p(x' - g, g)}{p(x - g, g)}\frac{m_h(x' - g)}{m_h(x - g)} \le \frac{p(x' - g', g')}{p(x - g', g')}\frac{m_h(x' - g')}{m_h(x - g')} = \frac{m_{h|1}(x' - g', g')}{m_{h|1}(x - g', g')}.$$

This inequality follows from two facts. First, due to the linearity of $V_0(\cdot)$,

$$\begin{aligned}\frac{p(x' - g, g)}{p(x - g, g)} &= \exp\left(\frac{V_1(x') - V_0(x' - g) - [V_1(x) - V_0(x - g)]}{\sigma}\right)\\ &= \exp\left(\frac{V_1(x') - V_0(x' - g') - [V_1(x) - V_0(x - g')]}{\sigma}\right) = \frac{p(x' - g', g')}{p(x - g', g')}.\end{aligned}$$

Second, consider the following inequality:

$$\frac{m_h(x' - g)}{m_h(x - g)} \le \frac{m_h(x' - g')}{m_h(x - g')}.$$

This holds if and only if $m_h(x' - g)/m_h(x - g)$ is increasing in g, that is,

$$\frac{\partial}{\partial g}\frac{m_h(x' - g)}{m_h(x - g)} = \frac{-m_h'(x' - g)m_h(x - g) + m_h(x' - g)m_h'(x - g)}{m_h(x - g)^2} \ge 0,$$

which is equivalent to

$$\frac{m_h'(x-g)}{m_h(x-g)} \geq \frac{m_h'(x'-g)}{m_h(x'-g)}.$$

This holds due to the log-concavity of $m_h(b)$ (Assumption 2). Therefore, for $g' > g$, $\tilde{m}_{h|1}(x, g')/\tilde{m}_{h|1}(x, g)$ increases with x when $x \geq g'$.

Since the support of $\tilde{m}_{h|1}(x, g)$ depends on g, I next establish the relationship between the monotone likelihood ratio dominance and first-order stochastic dominance with potentially different support.

Lemma B.10. *Consider a series of random variables indexed by i. For each i, the random variable is distributed over its support $[\underline{x}_i, \infty)$ with a cumulative distribution function $\Phi_i(x)$ and a probability density function $\phi_i(x)$. If $\underline{x}_i \leq \underline{x}_{i'}$ and $\phi_{i'}(x)/\phi_i(x)$ is increasing in x over $x \geq \underline{x}_{i'}$, then $\Phi_{i'}(x) \leq \Phi_i(x)$ for all x.*

Proof. For $x' > x \geq \underline{x}_{i'}$, we have the following monotone likelihood ratio dominance:

$$\frac{\phi_{i'}(x)}{\phi_{i'}(x')} \leq \frac{\phi_i(x)}{\phi_i(x')}. \tag{B.1}$$

Since $\phi_{i'}(x) = 0$ for $x < \underline{x}_{i'}$, (B.1) also holds for all $x \leq x'$. Therefore, integrating (B.1) over $x \leq x'$ gives

$$\int_{-\infty}^{x'} \frac{\phi_{i'}(x)}{\phi_{i'}(x')} dx \leq \int_{-\infty}^{x'} \frac{\phi_i(x)}{\phi_i(x')} dx,$$

which is equivalent to

$$-\frac{\phi_{i'}(x')}{\int_{-\infty}^{x'} \phi_{i'}(x)dx} \leq -\frac{\phi_i(x')}{\int_{-\infty}^{x'} \phi_i(x)dx}.$$

Integrating this over $x' \geq x''$, where $x'' \geq \underline{x}_{i'}$, gives

$$-\int_{x''}^{\infty} \frac{\phi_{i'}(x')}{\int_{-\infty}^{x'} \phi_{i'}(x)dx} dx' \leq -\int_{x''}^{\infty} \frac{\phi_i(x')}{\int_{-\infty}^{x'} \phi_i(x)dx} dx'. \tag{B.2}$$

Then, (B.2) can be written as follows: for all $x'' \geq \underline{x}_{i'}$,

$$\Phi_{i'}(x'') = \exp\left(-\int_{x''}^{\infty} \frac{\phi_{i'}(x')}{\int_{-\infty}^{x'} \phi_{i'}(x)dx} dx'\right) \leq \exp\left(-\int_{x''}^{\infty} \frac{\phi_i(x')}{\int_{-\infty}^{x'} \phi_i(x)dx} dx'\right) = \Phi_i(x'').$$

Moreover, for $x'' < \underline{x}_{i'}$, we have $\Phi_{i'}(x'') = 0 \leq \Phi_i(x'')$. □

This lemma implies that, for $g' > g$, the distribution of x conditional on g' first-order stochastically dominates that conditional on g. Since $V_1'(x)$ is decreasing in x (Lemma

4), $\int V_1'(b+g')m_{h|1}(b,g')db = \int V_1'(x)\tilde{m}_{h|1}(x,g')dx \leq \int V_1'(x)\tilde{m}_{h|1}(x,g)dx = \int V_1'(b+g)m_{h|1}(b,g)db$ follows.

B.12 Proof of Proposition 3

Part (i): Suppose that $m_{h'}(b)/m_h(b)$ is increasing in b for all b and $\hat{g}_{h'} > \hat{g}_h$. From the optimal policy formula (8), $\hat{g}_{h'} > \hat{g}_h$ implies $\int V_1'(b+\hat{g}_{h'})m_{h'|1}(b,\hat{g}_{h'})db > \int V_1'(b+\hat{g}_h)m_{h|1}(b,\hat{g}_h)db$. Moreover, since $m_{h'}(b)/m_h(b)$ is increasing in b , Lemma 5 implies $\int V_1'(b+\hat{g}_h)m_{h|1}(b,\hat{g}_h)db \geq \int V_1'(b+\hat{g}_h)m_{h'|1}(b,\hat{g}_h)db$. Therefore, we have $\int V_1'(b+\hat{g}_{h'})m_{h'|1}(b,\hat{g}_{h'})db > \int V_1'(b+\hat{g}_h)m_{h'|1}(b,\hat{g}_h)db$. By Lemma 7, this implies $\hat{g}_{h'} \leq \hat{g}_h$, leading to a contradiction. Therefore, $\hat{g}_{h'} \leq \hat{g}_h$ must hold.

Part (ii): Suppose that $m_h(b)/m_{h'}(b)$ is unimodal in b, $\int b\left[m_{h|1}(b,g)-m_{h'|1}(b,g)\right]db = 0$ for some g between $\hat{g}_h$ and $\hat{g}_{h'}$, $\gamma \geq 1/2$, and $\hat{g}_{h'} < \hat{g}_h$. From the optimal policy formula (8), $\hat{g}_{h'} < \hat{g}_h$ implies $\int V_1'(b+\hat{g}_{h'})m_{h'|1}(b,\hat{g}_{h'})db < \int V_1'(b+\hat{g}_h)m_{h|1}(b,\hat{g}_h)db$. Consequently, by Lemma 7, $\int V_1'(b+\hat{g})m_{h'|1}(b,\hat{g})db < \int V_1'(b+\hat{g})m_{h|1}(b,\hat{g})db$ also holds for $\hat{g} \in [\hat{g}_{h'}, \hat{g}_h]$ such that $\int b\left[m_{h'|1}(b,\hat{g}) - m_{h|1}(b,\hat{g})\right]db = 0$. However, since $m_h(b)/m_{h'}(b)$ is unimodal in b and $\gamma \geq 1/2$, Lemma 6 implies $\int V_1'(b+\hat{g})m_{h'|1}(b,\hat{g})db \geq \int V_1'(b+\hat{g})m_{h|1}(b,\hat{g})db$, leading to a contradiction. Therefore, $\hat{g}_{h'} \geq \hat{g}_h$ must hold.

B.13 Proof of Corollary 2

Under Assumption 3, $m_{h'}(b)/m_h(b) = \exp\left(b(1/\theta_h - 1/\theta_{h'})\right)\theta_h/\theta_{h'}$. Therefore, $m_{h'}(b)/m_h(b)$ is increasing in b if $\theta_{h'} \geq \theta_h$. Then, the result follows from Proposition 3(i).

B.14 Analytical Details on Section 3.4

Conditional Choice Probabilities and Values. Let $\boldsymbol{\varepsilon}_1 := (\varepsilon_j)_{j\in\mathcal{J}}$ denote a vector of psychic return for each college option j, distributed with a cumulative distribution function $F_1(\cdot)$. Then, the fraction of individuals choosing j, conditional on $(b,\boldsymbol{g})$ and attending college is

$$p_{j|1}(b,\boldsymbol{g}) := \int \mathbb{I}_{j\in\operatorname{argmax}_{j'\in\mathcal{J}}\{V_{j'}(b+g_{j'})+\varepsilon_{j'}\}} dF_1(\boldsymbol{\varepsilon}_1).$$

The average indirect utility of attending college conditional on $(b,\boldsymbol{g})$ is

$$V_1(b,\boldsymbol{g}) := \int \max_{j\in\mathcal{J}} \{V_j(b+g_j)+\varepsilon_j\} dF_1(\boldsymbol{\varepsilon}_1).$$

Finally, the average indirect utility conditional on $(b, \boldsymbol{g})$ is

$$V(b, \boldsymbol{g}) := \int \max\left\{V_0(b) + \varepsilon, V_1(b, \boldsymbol{g})\right\} dF(\varepsilon).$$

Efficiency Losses. The efficiency losses from intertemporal consumption distortion is

$$L_c(b, \boldsymbol{g}) := \sum_{j \in \mathcal{J}} p_j(b, \boldsymbol{g}) \left[(y_j + b + g_j - k_j) - V_j(b + g_j)\right].$$

The schooling distortion is

$$\begin{aligned} L_s(b, \boldsymbol{g}) := \int \max \left\{ y_0 + \varepsilon, \int \max_{j \in \mathcal{J}} \{y_j - k_j + \varepsilon_j\} dF_1(\boldsymbol{\varepsilon}) \right\} dF(\varepsilon) - \int \mathbb{I}_{\varepsilon > V_1(b, \boldsymbol{g}) - V_0(b)} (y_0 + \varepsilon) dF(\varepsilon) \\ - p(b, \boldsymbol{g}) \sum_{j \in \mathcal{J}} \int \mathbb{I}_{j \in \operatorname{argmax}_{j' \in \mathcal{J}} \{V_{j'}(b + g_{j'}) + \varepsilon_{j'}\}} (y_j - k_j + \varepsilon_j) dF_1(\boldsymbol{\varepsilon}_1). \end{aligned}$$

Lemma B.11. *The average indirect utility can be written as follows:*

$$V(b, \boldsymbol{g}) = \int \max \left\{ y_0 + \varepsilon, \int \max_{j \in \mathcal{J}} \{y_j - k_j + \varepsilon_j\} dF_1(\boldsymbol{\varepsilon}_1) \right\} dF(\varepsilon) + b + \sum_{j \in \mathcal{J}} p_j(b, \boldsymbol{g}) g_j - L(b, \boldsymbol{g}).$$

Proof. First, notice that we can rewrite $L_s(b, \boldsymbol{g})$ as follows:

$$\begin{aligned} L_s(b, \boldsymbol{g}) = & \int \max \left\{ y_0 + \varepsilon, \int \max_{j \in \mathcal{J}} \{y_j - k_j + \varepsilon_j\} dF_1(\boldsymbol{\varepsilon}_1) \right\} dF(\varepsilon) \\ & - [1 - p(b, \boldsymbol{g})] \left[y_0 - V_0(b)\right] - \sum_{j \in \mathcal{J}} p_j(b, \boldsymbol{g}) \left[y_j - k_j - V_j(b + g_j)\right] - V(b, \boldsymbol{g}), \end{aligned}$$

which follows from

$$\begin{aligned} & \int \mathbb{I}_{\varepsilon > V_1(b, \boldsymbol{g}) - V_0(b)} (y_0 + \varepsilon) dF(\varepsilon) + p(b, \boldsymbol{g}) \sum_{j \in \mathcal{J}} \int \mathbb{I}_{j \in \operatorname{argmax}_{j' \in \mathcal{J}} \{V_{j'}(b + g_{j'}) + \varepsilon_{j'}\}} (y_j - k_j + \varepsilon_j) dF_1(\boldsymbol{\varepsilon}_1) \\ = & \int \mathbb{I}_{\varepsilon > V_1(b, \boldsymbol{g}) - V_0(b)} \left[V_0(b) + \varepsilon + y_0 - V_0(b)\right] dF(\varepsilon) \\ & + p(b, \boldsymbol{g}) \sum_{j \in \mathcal{J}} \int \mathbb{I}_{j \in \operatorname{argmax}_{j' \in \mathcal{J}} \{V_{j'}(b + g_{j'}) + \varepsilon_{j'}\}} \left[V_j(b + g_j) + \varepsilon_j + y_j - k_j - V_j(b + g_j)\right] dF_1(\boldsymbol{\varepsilon}_1) \\ = & [1 - p(b, \boldsymbol{g})] \left[y_0 - V_0(b)\right] + \sum_{j \in \mathcal{J}} p_j(b, \boldsymbol{g}) \left[y_j - k_j - V_j(b + g_j)\right] - V(b, \boldsymbol{g}). \end{aligned}$$

Therefore,

$$\begin{aligned} L_c(b, \boldsymbol{g}) + L_s(b, \boldsymbol{g}) = & \int \max \left\{ y_0 + \varepsilon, \int \max_{j \in \mathcal{J}} \{y_j - k_j + \varepsilon_j\} dF_1(\boldsymbol{\varepsilon}_1) \right\} dF(\varepsilon) + b \\ & + \sum_{j \in \mathcal{J}} p_j(b, \boldsymbol{g}) g_j - V(b, \boldsymbol{g}). \end{aligned}$$

□

Planning Problem. Define

$$P_{j,h}(\boldsymbol{g}) := \int p_j(b,\boldsymbol{g}) m_h(b) db.$$

Definition 2. *The optimal policy $(\hat{\boldsymbol{g}}_h)_{h\in\mathcal{H}}$ solves*

$$\min_{(\boldsymbol{g}_h)_{h\in\mathcal{H}}} \sum_{h\in\mathcal{H}} n_h \int L(b,\boldsymbol{g}_h) m_h(b) db$$

$$\text{subject to } \sum_{h\in\mathcal{H}} n_h \sum_{j\in\mathcal{J}} P_{j,h}(b,\boldsymbol{g}_h) g_{j,h} \le G, \tag{B.3}$$

and $g_{j,h} \ge 0$ for all $(j,h) \in \mathcal{J}\times\mathcal{H}$. Let $\lambda \ge 0$ be the Lagrange multiplier on the constraint (B.3).

The first order condition with respect to $g_{j,h}$ is

$$\int V_j'(b+\hat{g}_{j,h}) m_{h|j}(b,\hat{\boldsymbol{g}}_h) db \le (1+\lambda)\left[1 + \sum_{j'\in\mathcal{J}} \frac{\partial P_{h,j'}(\hat{\boldsymbol{g}}_h)}{\partial g_{j,h}} \frac{\hat{g}_{j',h}}{P_{j,h}(\hat{\boldsymbol{g}}_h)}\right].$$

Proposition B.7. *Suppose Assumptions 1 and 4 hold. If $\hat{g}_{j,h} > 0$ and $p(b,\hat{\boldsymbol{g}}_h) < 1$ for all b, then $\hat{g}_{j,h}$ satisfies* (9).

Proof. With Assumption 4,

$$p_{j|1}(b,\boldsymbol{g}) = \frac{\exp\left(\left[V_j(b+g_j)+\mu_j\right]/\sigma\right)}{\sum_{j'\in\mathcal{J}} \exp\left(\left[V_{j'}(b+g_{j'})+\mu_{j'}\right]/\sigma\right)},$$

$$V_1(b,\boldsymbol{g}) = \sigma \ln\left(\sum_{j\in\mathcal{J}} \exp\left(\left[V_j(b+g_j)+\mu_j\right]/\sigma\right)\right).$$

Then, with Assumption 1, when $p(b,\boldsymbol{g}) < 1$,

$$p(b,\boldsymbol{g}) = \exp\left(\frac{V_1(b,\boldsymbol{g}) - V_0(b) - \mu - \sigma}{\sigma}\right) = \exp\left(\frac{-V_0(b)-\mu-\sigma}{\sigma}\right)\sum_{j\in\mathcal{J}} \exp\left(\left[V_j(b+g_j)+\mu_j\right]/\sigma\right).$$

Therefore,

$$p_j(b,\boldsymbol{g}) = p(b,\boldsymbol{g}) p_{j|1}(b,\boldsymbol{g}) = \exp\left(\frac{V_j(b+g_j)+\mu_j - V_0(b) - \mu - \sigma}{\sigma}\right).$$

This implies

$$\frac{\partial P_{j',h}(\boldsymbol{g})}{\partial g_j}\frac{1}{P_{j,h}(\boldsymbol{g})} = \begin{cases} \frac{1}{\sigma}\int V_j'(b+g_j)m_{h|j}(b,\boldsymbol{g})db, & \text{for } j'=j, \\ 0, & \text{for } j' \neq j. \end{cases}$$

Therefore, the first order condition for the case of $\hat{g}_{j,h} > 0$ and $p(b, \hat{\boldsymbol{g}}_h) < 1$ is

$$\int V_j'(b+\hat{g}_{j,h})m_{h|j}(b,\hat{\boldsymbol{g}}_h)db = (1+\lambda)\left[1+\frac{\hat{g}_{j,h}}{\sigma}\int V_j'(b+\hat{g}_{j,h})m_{h|j}(b,\hat{\boldsymbol{g}}_h)db\right].$$

Rearranging this equation gives (9). □

Lemma B.12. *Suppose that Assumptions 1, 2, and 4 hold. If $p(b,\boldsymbol{g}) < 1$ for all b, then $\int V_j(b+g_j)m_{h|j}(b,\boldsymbol{g})db$ is decreasing in g_j.*

Proof. It is straightforward to see that the proof of Lemma 7 also applies to this case, once we replace $V_1(b+g)$ with $V_j(b+g_j)$. □

B.15 Proof of Proposition 4

It is straightforward to see that the proof of Proposition 3 also applies to this case, once we replace $V_1(b+g)$ with $V_j(b+g_j)$.

B.16 Proof of Lemma 8

Since $g_{j'} = g_j$ and $R(y_j-\bar{l})/(b+g_j-k_j+\bar{l})$ is increasing in k_j and y_j, we have

$$\frac{V_{j'}'(b+g_{j'})}{V_j'(b+g_j)} = \left[\frac{1+\beta\max\left\{1, R(y_{j'}-\bar{l})/(b+g_{j'}-k_{j'}+\bar{l})\right\}^{\frac{\gamma-1}{\gamma}}}{1+\beta\max\left\{1, R(y_j-\bar{l})/(b+g_j-k_j+\bar{l})\right\}^{\frac{\gamma-1}{\gamma}}}\right]^{\frac{1}{\gamma-1}} \geq 1. \tag{B.4}$$

Next, $m_{h|j'}(b,\boldsymbol{g})/m_{h|j}(b,\boldsymbol{g})$ is increasing in b due to (B.4):

$$\frac{m_{h|j'}(b,\boldsymbol{g})}{m_{h|j}(b,\boldsymbol{g})} = \frac{p_{j'}(b,\boldsymbol{g})}{p_j(b,\boldsymbol{g})}\frac{m_h(b)}{m_h(b)}\frac{P_{h,j}(\boldsymbol{g})}{P_{h,j'}(\boldsymbol{g})} = \exp\left(\frac{V_{j'}(b+g_{j'})+\mu_{j'}-V_j(b+g_j)-\mu_j}{\sigma}\right)\frac{P_{h,j}(\boldsymbol{g})}{P_{h,j'}(\boldsymbol{g})}.$$

B.17 Proof of Proposition 5

Part (i): It is straightforward to see that Corollary 2 applies to this case, once we replace $V_1(b+g)$ with $V_j(b+g_j)$.

Part (ii): Define

$$\omega(x) := \left(\frac{1+\beta\max\{1,x^{-1}\}^{\frac{\gamma-1}{\gamma}}}{1+\beta}\right)^{\frac{1}{\gamma-1}} \geq 1,$$

which is decreasing in x. Then, by a change of variables $x := (b + g_j - k_j + \bar{l})/[R(y_j - \bar{l})]$,

$$V_j'(b + g_j) = \omega\left(\frac{b + g_j - k_j + \bar{l}}{R(y_j - \bar{l})}\right) \Leftrightarrow V_j'\big(xR(y_j - \bar{l}) + k_j - \bar{l}\big) = \omega(x).$$

The density of x conditional on choice j is $m_{h|j}\big(xR(y_j - \bar{l}) - g_j + k_j - \bar{l}, \boldsymbol{g}\big)R(y_j - \bar{l})$, so

$$\int_0^\infty V_j'(b + g_j)m_{h|j}(b, \boldsymbol{g})db = \int_{\frac{g_j - k_j + \bar{l}}{R(y_j - \bar{l})}}^\infty \omega(x)m_{h|j}\big(xR(y_j - \bar{l}) - g_j + k_j - \bar{l}, \boldsymbol{g}\big)R(y_j - \bar{l})dx.$$

For $j' > j$, consider the ratio between conditional densities of x, where these densities are strictly positive ($x \geq \max\{(g_j - k_j + \bar{l})/[R(y_j - \bar{l})], (g_{j'} - k_{j'} + \bar{l})/[R(y_{j'} - \bar{l})]\}$):

$$\begin{aligned}
&\frac{m_{h|j'}\big(xR(y_{j'} - \bar{l}) - g_{j'} + k_{j'} - \bar{l}, \boldsymbol{g}\big)}{m_{h|j}\big(xR(y_j - \bar{l}) - g_j + k_j - \bar{l}, \boldsymbol{g}\big)} \frac{(y_{j'} - \bar{l})}{(y_j - \bar{l})} \\
&= \exp\left(\frac{V_{j'}\big(xR(y_{j'} - \bar{l}) + k_{j'} - \bar{l}\big) + \mu_{j'} - V_j\big(xR(y_j - \bar{l}) + k_j - \bar{l}\big) - \mu_j}{\sigma}\right) \\
&\quad \times \frac{m_h\big(xR(y_{j'} - \bar{l}) - g_{j'} + k_{j'} - \bar{l}\big)}{m_h\big(xR(y_j - \bar{l}) - g_j + k_j - \bar{l}\big)} \frac{P_{j,h}(\boldsymbol{g})}{P_{j',h}(\boldsymbol{g})} \frac{(y_{j'} - \bar{l})}{(y_j - \bar{l})}.
\end{aligned}$$

Taking a derivative with respect to x after taking logs gives

$$\frac{\omega(x)}{\sigma}R(y_{j'} - y_j) + \frac{m_h'\big(xR(y_{j'} - \bar{l}) - g_j + k_{j'} - \bar{l}\big)}{m_h\big(xR(y_{j'} - \bar{l}) - g_{j'} + k_{j'} - \bar{l}\big)}R(y_{j'} - \bar{l}) - \frac{m_h'\big(xR(y_j - \bar{l}) - g_j + k_j - \bar{l}\big)}{m_h\big(xR(y_j - \bar{l}) - g_j + k_j - \bar{l}\big)}R(y_j - \bar{l}).$$

Under Assumption 3, this is simplified to

$$\left[\frac{\omega(x)}{\sigma} - \frac{1}{\theta_h}\right]R(y_{j'} - y_j). \tag{B.5}$$

Note that, due to $b \geq 0$, $g_j \geq 0$, and the monotonicity of y_j and k_j in j,

$$x = \frac{b + g_j - k_j + \bar{l}}{R(y_j - \bar{l})} \geq \frac{\bar{l} - k_j}{R(y_j - \bar{l})} \geq \frac{\bar{l} - k_J}{R(y_J - \bar{l})}.$$

Since $\omega(x)$ is decreasing in x,

$$\omega(x) \leq \omega\left(\frac{\bar{l} - k_J}{R(y_J - \bar{l})}\right) = \max_{j \in \mathcal{J}} V_j'(0).$$

Therefore, (B.5) is negative when $\theta_h \leq \sigma/\max_{j\in\mathcal{J}} V_j'(0)$. In this case, by Lemma B.10, the distribution of x conditional on j first-order stochastically dominates the distribution of x

conditional on j' for $j' > j$ with $g_{j'} \le g_j$. Since $\omega(x)$ is decreasing in x, this implies

$$\begin{aligned}&\int_{\frac{g_{j'}-k_{j'}+\bar{l}}{R(y_{j'}-\bar{l})}}^{\infty} \omega(x) m_{h|j'}\big(xR(y_{j'}-\bar{l}) - g_{j'} + k_{j'} - \bar{l}, \boldsymbol{g}\big) R(y_{j'}-\bar{l}) dx \\ &\ge \int_{\frac{g_j-k_j+\bar{l}}{R(y_j-\bar{l})}}^{\infty} \omega(x) m_{h|j}\big(xR(y_j-\bar{l}) - g_j + k_j - \bar{l}, \boldsymbol{g}\big) R(y_j-\bar{l}) dx,\end{aligned}$$

which is equivalent to

$$\int_0^{\infty} V'_{j'}(b+g_{j'}) m_{h|j'}(b,\boldsymbol{g}) db \ge \int_0^{\infty} V'_j(b+g_j) m_{h|j}(b,\boldsymbol{g}) db. \tag{B.6}$$

Suppose that $\theta_h \le \sigma / \max_{j\in\mathcal{J}} V'_j(0)$ and $\hat{g}_{j',h} < \hat{g}_{j,h}$ for $j' > j$. Then, the formula (9) implies $\int_0^{\infty} V'_{j'}(b+\hat{g}_{j',h}) m_{h|j'}(b,\hat{\boldsymbol{g}}) db < \int_0^{\infty} V'_j(b+\hat{g}_{j,h}) m_{h|j}(b,\hat{\boldsymbol{g}}) db$. This contradicts (B.6), thus $\hat{g}_{j',h} \ge \hat{g}_{j,h}$ must hold.

B.18 Proof of Lemma 9

When (12) does not bind, $V_j(b_j+g_j;a) = y_j(a) + b_j + g_j - k_j$ and $V'_j(b_j+g_j;a) = 1$. When (12) binds,

$$V_j(b_j+g_j;a) = \Lambda^{\frac{1}{1-\gamma}} \left(\sum_{t=1}^{T_j} \beta^t c_{k,1}^{\frac{\gamma-1}{\gamma}} + \sum_{t=T_j+1}^{T_k} \beta^t c_{k,2}^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}},$$

where

$$c_{k,1} := \frac{b_j + g_j + \bar{l}_j - k_j}{\sum_{t=1}^{T_j} R^{-t}}, \quad c_{k,2} := \frac{y_j(a) - \bar{l}_j}{\sum_{t=T_j+1}^{T_k} R^{-t}}.$$

Therefore,

$$V'_j(b_j+g_j;a) = \Lambda^{\frac{1}{1-\gamma}} \left(\sum_{t=1}^{T_j} \beta^t c_{k,1}^{\frac{\gamma-1}{\gamma}} + \sum_{t=T_j+1}^{T_k} \beta^t c_{k,2}^{\frac{\gamma-1}{\gamma}} \right)^{\frac{1}{\gamma-1}} c_{k,1}^{\frac{-1}{\gamma}} = \left[\frac{\sum_{t=1}^{T_j} \beta^t}{\sum_{t=1}^{T_k} \beta^t} + \frac{\sum_{t=T_j+1}^{T_k} \beta^t}{\sum_{t=1}^{T_k} \beta^t} \left(\frac{c_{k,2}}{c_{k,1}} \right)^{\frac{\gamma-1}{\gamma}} \right]^{\frac{1}{\gamma-1}},$$

which is greater than 1 because $c_{k,2} \ge c_{k,1}$, that is, the annual consumption during schooling is lower than the annual consumption after schooling when (12) binds.

Next, I show that $V'_j(b_j+g_j;a)$ is increasing in a. When (12) does not bind, the amount of borrowing can be written as

$$\sum_{t=1}^{T_j} R^{-t} c_t + k_j - b_j - g_j = \sum_{t=1}^{T_j} R^{-t} \frac{y_j(a) + b_j + g_j - k_j}{\sum_{t=1}^{T_k} R^{-t}} + k_j - b_j - g_j,$$

which is increasing in a. Therefore, (12) binds for individuals with higher values of a. When (12) binds, $V_j'(b_j + g_j; a)$ increases with a, as $c_{k,2}/c_{k,1}$ is increasing in a.

B.19 Proof of Proposition 6

With the distributional assumption, the average value of attending college and the conditional choice probability are written as follows:

$$V_1(\boldsymbol{b}, \boldsymbol{g}, a, h) = \tilde{\sigma}_1 \ln\left(\sum_{j\in\mathcal{J}} \exp\left(\left[V_j(b_j + g_j; a, h) + \mu_j(a, h)\right]/\tilde{\sigma}_1\right)\right),$$

$$p_{j|1}(\boldsymbol{b}, \boldsymbol{g}, a, h) = \frac{\exp\left(\left[V_j(b_j + g_j; a, h) + \mu_j(a, h)\right]/\tilde{\sigma}_1\right)}{\sum_{j'\in\mathcal{J}} \exp\left(\left[V_{j'}(b_{j'} + g_{j'}; a, h) + \mu_{j'}(a, h)\right]/\tilde{\sigma}_1\right)},$$

With the logistic distribution, the average indirect utility and probability of college attendance are given by

$$V(\boldsymbol{b}, \boldsymbol{g}, a, h) = \tilde{\sigma} \ln\left(\exp\left(V_0(b_0; a, h)/\tilde{\sigma}\right) + \exp\left(V_1(\boldsymbol{b}, \boldsymbol{g}, a, h)/\tilde{\sigma}\right)\right),$$

$$p(\boldsymbol{b}, \boldsymbol{g}, a, h) = \frac{\exp\left(V_1(\boldsymbol{b}, \boldsymbol{g}, a, h)/\tilde{\sigma}\right)}{\exp\left(V_0(b_0; a, h)/\tilde{\sigma}\right) + \exp\left(V_1(\boldsymbol{b}, \boldsymbol{g}, a, h)/\tilde{\sigma}\right)},$$

The derivative of $V(\boldsymbol{b}, \boldsymbol{g}, a, h)$ with respect to b_j is

$$\frac{\partial V(\boldsymbol{b}, \boldsymbol{g}, a, h)}{\partial b_j} = p_j(\boldsymbol{b}, \boldsymbol{g}, a, h) V_j'(b_j + g_j; a).$$

Therefore, the first order condition for an interior solution is

$$\frac{1}{p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)} \frac{\partial}{\partial b_j}\left(\sum_{j'\in\overline{\mathcal{J}}} p_{j'}(\boldsymbol{b}, \boldsymbol{g}, a, h) b_{j'}\right) = \left(\frac{\delta}{1-\delta}\right) \frac{v'(\hat{c}_k)}{v'(\hat{c}_p)} V_j'(\hat{b}_j + g_j; a). \tag{B.7}$$

Part (ii): First, note that $\tilde{\sigma} = \tilde{\sigma}_1$ implies

$$p_j(\boldsymbol{b}, \boldsymbol{g}, a, h) = \frac{\exp\left(\left[V_j(b_j + g_j; a) + \mu_j(a, h)\right]/\tilde{\sigma}\right)}{\sum_{j'\in\overline{\mathcal{J}}} \exp\left(\left[V_{j'}(b_j + g_j; a) + \mu_j(a, h)\right]/\tilde{\sigma}\right)}, \quad \forall j \in \overline{\mathcal{J}},$$

$$V(\boldsymbol{b}, \boldsymbol{g}, a, h) = \tilde{\sigma} \ln\left(\sum_{j\in\overline{\mathcal{J}}} \exp\left(\left[V_j(b_j + g_j; a) + \mu_j(a, h)\right]/\tilde{\sigma}\right)\right),$$

where we define $g_0 = \mu_0(a, h) = 0$.

By differentiating $p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)$ with respect to b_j, we get

$$\frac{\partial p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)}{\partial b_j} = p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)\left[1 - p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)\right]\frac{V_j'(b_j + g_j; a)}{\tilde{\sigma}}.$$

The derivative of $p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)$ with respect to $b_{j'}$ for $j' \neq j$ is

$$\frac{\partial p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)}{\partial b_{j'}} = -p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)p_{j'}(\boldsymbol{b}, \boldsymbol{g}, a, h)\frac{V_{j'}'(b_{j'} + g_{j'}; a)}{\tilde{\sigma}}.$$

The derivative of average parental transfer with respect to b_j is

$$\frac{\partial}{\partial b_j}\left(\sum_{j' \in \overline{\mathcal{J}}} p_{j'}(\boldsymbol{b}, \boldsymbol{g}, a, h)b_{j'}\right) = p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)\left\{1 + \left(b_j - \sum_{j' \in \overline{\mathcal{J}}} p_{j'}(\boldsymbol{b}, \boldsymbol{g}, a, h)b_{j'}\right)\frac{V_j'(b_j + g_j; a)}{\tilde{\sigma}}\right\}.$$

Substituting the above condition into (B.7) and rearranging gives (16).

Part (i): When $\tilde{\sigma} \neq \tilde{\sigma}_1$, the derivative of average parental transfer with respect to b_j is more complicated.

For $j \in \mathcal{J}$ such that $V_j'(b_j + g_j; a) = 1$,

$$\begin{aligned}&\frac{1}{p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)}\frac{\partial}{\partial b_j}\left(\sum_{j' \in \overline{\mathcal{J}}} p_{j'}(\boldsymbol{b}, \boldsymbol{g}, a, h)b_{j'}\right)\\ &= \left\{1 + \frac{b_j}{\sigma_1} - \frac{p_0(\boldsymbol{b}, \boldsymbol{g}, a, h)b_0}{\sigma} + \left[\frac{p_0(\boldsymbol{b}, \boldsymbol{g}, a, h)}{\sigma} - \frac{1}{\sigma_1}\right]\sum_{j' \in \mathcal{J}} p_{j'|1}(\boldsymbol{b}, \boldsymbol{g}, a, h)b_{j'}\right\}.\end{aligned}$$

Substituting the condition above into (B.7) and taking a difference between $j \in \mathcal{J}$ and $j' \in \mathcal{J}$ implies $b_j = b_{j'}$ for $j \in \mathcal{J}$ and $j' \in \mathcal{J}$.

Similarly, for $j = 0$,

$$\frac{1}{p_0(\boldsymbol{b}, \boldsymbol{g}, a, h)}\frac{\partial}{\partial b_0}\left(\sum_{j' \in \overline{\mathcal{J}}} p_{j'}(\boldsymbol{b}, \boldsymbol{g}, a, h)b_{j'}\right) = \left\{1 + \frac{1}{\sigma}\left[b_0 - \sum_{j' \in \overline{\mathcal{J}}} p_{j'}(\boldsymbol{b}, \boldsymbol{g}, a, h)b_{j'}\right]\right\}.$$

Differencing the first order condition between $j \in \mathcal{J}$ and $j' = 0$ leads to

$$\begin{aligned}&\frac{1}{p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)}\frac{\partial}{\partial b_j}\left(\sum_{j' \in \overline{\mathcal{J}}} p_{j'}(\boldsymbol{b}, \boldsymbol{g}, a, h)b_{j'}\right) - \frac{1}{p_0(\boldsymbol{b}, \boldsymbol{g}, a, h)}\frac{\partial}{\partial b_0}\left(\sum_{j' \in \overline{\mathcal{J}}} p_{j'}(\boldsymbol{b}, \boldsymbol{g}, a, h)b_{j'}\right)\\ &= \frac{b_j - b_0}{\sigma} = 0.\end{aligned}$$

Therefore, $b_j = b_{j'}$ for all $j \in \overline{\mathcal{J}}$ and $j' \in \overline{\mathcal{J}}$.

C Details on Parameterization

C.1 Monetary Returns to, and Costs of, Schooling

Table C.1: Determinants of Log Annual Earnings

	Coefficient	Standard Error
Schooling Choice:		
2-Year Only	0.365	0.017
2-Year and 4-Year	0.486	0.023
4-Year Only	0.565	0.013
AFQT Quartile:		
2	0.286	0.014
3	0.376	0.015
4	0.441	0.017
Potential Experience:		
1	0.427	0.025
2	0.722	0.025
3	0.871	0.025
4	1.009	0.025
5	1.089	0.025
6	1.130	0.025
7	1.139	0.026
8	1.192	0.026
9	1.218	0.029
10	1.287	0.028
11	1.322	0.031
12	1.341	0.029
13	1.417	0.033
14	1.406	0.032
15	1.427	0.037
16	1.473	0.035
Constant	8.278	0.020
Observations	34,757	
R-Squared	0.244	

Notes: OLS estimates based on before-tax real earnings.

Table C.2: Present Discounted Value of Lifetime Earnings at Age 17 ($)

Schooling Choice	$y_j(a)$ for AFQT Quartiles			
	1	2	3	4
No College	268,686	357,538	391,070	417,244
2-Year Only	356,074	473,910	518,380	553,092
2- and 4-Year	369,337	491,585	537,720	573,731
4-Year Only	399,720	532,039	581,975	620,953

Table C.3: Average Annual Amounts for Each Institution ($)

Institution	Tuition	Books	Earnings	Grants for Income Quartiles				Staffod Loan Limit for Years		
				1	2	3	4	1	2	3+
2-Year	1,886	1,038	5,032	3,702	1,904	977	762	2,625	3,500	
4-Year	4,769		3,112	6,526	3,518	1,738	1,502			5,500

Table C.4: Present Discounted Values at Age 17 ($)

Schooling Choice	Cost (k_j)	Aid for Income Quartiles ($g_{j,h}$)				Borrowing Limit ($\bar{l}_j$)
		1	2	3	4	
2-Year Only	-4,033	7,084	3,642	1,870	1,458	5,848
2- and 4-Year	11,336	22,910	14,875	10,343	9,085	15,768
4-Year Only	31,673	32,615	23,148	17,461	15,719	15,768

Notes: Cost is calculated as tuition plus books, minus earnings while enrolled.

C.2 Parental Transfers from the NLSY97 Data

The NLSY97 survey contains two sections in which questions about parental transfers are asked. The "college experience" section collects information about how much money youth received from parents to pay for college and the "income" section asks about the money youth received as part of their income in each year.

C.2.1 College Experience Section

In every interview, the college experience section starts with question YSCH-24991, which asks if the respondent attended college since the date of last interview (DLI). For each college and each term the respondent attended since the DLI, the survey asks, among other things, for the start date of the term (YSCH-20400) and how the respondent paid for college. The structure of questions on college financing differs depending on whether it is the first time the respondent is asked about the college in each round (YSCH-22004). For the first term of each college in each round, the respondent is first asked whether she received financial aid from biological parents, mother (and stepfather), father (and stepmother), grandparents, and other relatives and friends (YSCH-23900) and, if so, the amount of financial aid received from each of them in the form of gift (YSCH-24600) and loan (YSCH-24700). The exact wording of YSCH-24600 is as follows:

> Altogether, how much [have/has/did] your family and friends [give/given] you in gifts or other money you are not expected to repay to help pay for your attendance at this school/institution during this term?

and the wording for YSCH-24700 is as follows:

> Altogether, how much [have/has/did] your family and friends [loan/loaned] you to help pay for your attendance at this school/institution during this term?

From the second term for each college in each round, the respondent is asked whether there were any changes in college financing since the previous term (YSCH-22005), and no further questions are asked if there were no changes. In this case, I fill in this information using answers from previous terms. If there were changes, then the respondent is first asked whether she received any financial assistance from family (YSCH-22006), and if she did, the further questions on aid from family (YSCH-23900, YSCH-24600, and YSCH-24700) are asked.

In round 1, the structure of survey is a little bit different from what is described above. I do not use data from the first round of the survey because financial aid questions are not asked for each term. This is likely to have very little effect because only eight respondents attended college during the first round of the survey, which was administered in 1997.

C.2.2 Income Section

The income section of the survey collects wage and salary data for the past calendar year from all respondents. Those who are considered independent answer more extensive questions about other sources of income, including parental transfers. Since one of the criteria for independence is reaching the age of 18, I measure parental transfers youth received when they are more than or equal to 18 years old.[5]

For rounds 1–7 (survey years 1997–2003), respondents state the amount of money they received from (i) both parent figures, (ii) the mother figure, and (iii) the father figure. Youth who live with both parent figures (YINC-5600) are asked whether they received any money from them in the previous year (YINC-5700), and if so, how much (YINC-5800). In round 2, the exact wording of YINC-5700 is as follows:

> Other than an allowance, did your parents give [you/you or your spouse] any money during 1997? Please include any gifts in the form of cash or a check but do not include any loans from your parents.

The wording for YINC-5800 is as follows:

> How much did your parents give [you/you and your spouse] during 1997?

Those who do not know the answer or refuse to answer question YINC-5800 are again asked to provided answers in categories (YINC-5900).

Those who live with a mother/father figure or whose biological mother/father is alive (YINC-6400/YINC-7000), including those who live with both parent figures, are also asked similar questions about transfers received from their mother figure (YINC-6500, YINC-6600, and YINC-6700) and father figure (YINC-7100, YINC-7200, and YINC-7300).

For rounds 8–15 (survey years 2004–2011), youth are first asked whether they received money from family or friends (YINC-5700A), and if so, their amounts in categories (YINC-5900A). These questions are discontinued after round 15. Therefore, parental transfers can be measured only through calendar year 2010.

Finally, for all rounds, youth are asked whether they received money from inheritances (YINC-5200), and if so, their amounts in numbers (YINC-5300) or categories (YINC-5400). In round 2, the exact wording of YINC-5200 is as follows:

> During 1997, did [you/you or your spouse/partner] receive any property or money from any estates, trusts, annuities or inheritances?

And the wording of YINC-5300 is as follows:

> What was the total market value or amount that [you/you and your spouse/partner] received during 1997 from these sources?

[5] Youth are also considered to be independent if they have had a child, are enrolled in a four-year college, are no longer enrolled in school, are not living with any parents or parent-figures, or have ever been married or are in a marriage-like relationship at the time of the survey.

C.2.3 Total Parental Transfers

For each calendar year, I construct annual parental transfers by adding all forms of college financial aid from all family members (YSCH-24600 and YSCH-24700) during all terms that started in that year to the amount of money received from family (YINC-5800/5900, YINC-6600/6700, YINC-7200/7300, and YINC-5900A) and inheritances (YINC-5300/5400). The variables provided in categories (YINC-5900, YINC-6700, YINC-7300, YINC-5400, and YINC-5900A) are turned into amounts by using the mid-point of each category or the minimum value for the highest category. Based on the annual parental transfers, I compute total parental transfers youth received between the years 1998 and 2010 and the ages 18 and 26, discounted back to age 17.

There are several issues related to constructing the total parental transfer variable. First, as discussed by Abbott et al. (2019), it might be incorrect to sum all transfers from different sections of the survey, because they are not necessarily mutually exclusive categories. For example, respondents might consider the college financial aid from family (recorded in the college experience section) a part of their income received from family (reported in the income section). Second, I include both family loans (YSCH-24700) and gifts, since the distinction between the two is unlikely to have meaningful policy implications—as long as parental support helps mitigate distortions from borrowing constraints. Third, unlike Johnson (2013) and Abbott et al. (2019), I do not include the monetary value of living with parents to parental transfers because it does not necessarily reflect parental support that helps youth attend college. Of course, students living at home while enrolled can save considerably on room and board costs, but this is useful only if there is a college that youth wish to attend within commuting distance from home.[6] Moreover, parents who are not willing give money for college might still let their children to stay with them because the additional cost of doing so would be small as a result of economies of scale. Therefore, it is possible that youth who do not live near a college stay at home because they do not have enough money to leave home and attend college elsewhere. Another reason for not including parental co-residence is that obtaining reliable information about it is difficult after round 6 because, as noted by Kaplan (2012), one must rely on the household roster that may refer to what youth consider their primary residence rather than their current residence. For example, college students who live away from home during the school year may still report to be living at home.

C.3 Identification

I first discuss identification of the distribution of measurement errors in parental transfers as well as the distribution of true parental transfers. In addition to the conditional independence

[6] Do (2004) notes that about half of the high school sophomore class of 1980 in the US did not have a public university in their county of residence.

assumption (19) and the normalization that $\Pr(B \in \mathcal{B}_1 | B^* \in \mathcal{B}_{k^*})$ decreases in k^*, I make two assumptions required for identification. For a given value of $(J, \boldsymbol{X})$, consider a $K \times Q$ matrix with elements $\Pr(B \in \mathcal{B}_k, \boldsymbol{Z} = z_q | J, \boldsymbol{X})$. I assume that the rank of this matrix is greater than or equal to K. Moreover, I assume that $\mathrm{E}[\tilde{B} | B^* \in \mathcal{B}_{k^*}, J, \boldsymbol{X}]$ takes distinct values for different values of k^*. Then, by a discrete-distribution version of Hu and Schennach (2008), as presented by Hu (2008), the joint distribution $\Pr(B \in \mathcal{B}_k, \tilde{B} \in \tilde{B}_l, \boldsymbol{Z}, B^* \in \mathcal{B}_{k^*} | J, \boldsymbol{X})$ is identified from the joint distribution $\Pr(B \in \mathcal{B}_k, \tilde{B} \in \tilde{B}_l, \boldsymbol{Z} | J, \boldsymbol{X})$ that takes the following form due to the conditional independence assumption (19):

$$\sum_{k^*=1}^{K} \Pr(B \in \mathcal{B}_k | B^* \in \mathcal{B}_{k^*}) \Pr(\tilde{B} \in \tilde{\mathcal{B}}_l | B^* \in \mathcal{B}_{k^*}) \Pr(\boldsymbol{Z} | \boldsymbol{X}, J, B^* \in \mathcal{B}_{k^*}) \Pr(B^* \in \mathcal{B}_{k^*} | \boldsymbol{X}, J).$$

Therefore, the parameters for the conditional distribution of measured parental transfers—ζ_{k,k^*} and $\xi_{l,k}$—as well as the conditional distribution of parental transfers $\Pr(B^* \in \mathcal{B}_{k^*} | \boldsymbol{X}, \boldsymbol{Z}, J)$ are identified.

Next, the parameters related to the psychic return to schooling and are identified from schooling distribution of unconstrained youth who receive large amount of parental transfers. Since $\Pr(B \in \mathcal{B}_k, \tilde{B} \in \tilde{B}_l, \boldsymbol{Z}, B^* \in \mathcal{B}_{k^*} | J, \boldsymbol{X})$ is identified, the joint distribution $\Pr(J | B^* \in \mathcal{B}_{k^*}, \boldsymbol{X})$ is known. Suppose that there exists k^* such that none of the constraints (12) binds for those with $B^* \in \mathcal{B}_{k^*}$ and $\boldsymbol{X}$. As shown in Proposition 6, parental transfer amounts are identical across schooling choices in this case. Therefore, parental transfer does not affect schooling choice. Let $\boldsymbol{b} = (b_j)_{j \in \overline{\mathcal{J}}}$, where $b_j = b \in \mathcal{B}_k^*$ for all $j \in \overline{\mathcal{J}}$.

For j and j' in $\mathcal{J}$,

$$\ln\left(\frac{p_j(\boldsymbol{b}, \boldsymbol{g}_h, a, h)}{p_{j'}(\boldsymbol{b}, \boldsymbol{g}_h, a, h)}\right) = \frac{\left[y_j(a) + g_{j,h} - k_j + \psi_j\right] - \left[y_{j'}(a) + g_{j',h} - k_{j'} + \psi_{j'}\right]}{\tilde{\sigma}_1}, \tag{C.1}$$

where only $\psi_j - \psi_{j'}$ and $\tilde{\sigma}_1$ are unknown. Since $y_j(a)$ is already known, single-differencing (C.1) in a (i.e., between two groups a and a', holding everything else constant) identifies $\tilde{\sigma}_1$. Then $\psi_j - \psi_{j'}$ is identified from (C.1).

Similarly, the log odds ratio for the extensive margin choice is

$$\ln\left(\frac{p(\boldsymbol{b}, \boldsymbol{g}_h, a, h)}{1 - p(\boldsymbol{b}, \boldsymbol{g}_h, a, h)}\right) = \frac{\psi_j + \kappa_a + \nu_h + \Gamma_{a,h}}{\tilde{\sigma}}, \tag{C.2}$$

where $\Gamma_{a,h} := \tilde{\sigma}_1 \ln \sum_{j' \in \mathcal{J}} \exp\left(\left[y_{j'}(a) + g_{j',h} - k_{j'} + (\psi_{j'} - \psi_j)\right] / \tilde{\sigma}_1\right) - y_0(a)$ is already identified above. Double-differencing (C.2) in (a, h) identifies $\tilde{\sigma}$. Since $\kappa_a = \nu_h = 0$ for some (a, h), single-differencing (C.2) in a identifies κ_a, while ν_h is identified by single-differencing in h. Then ψ_j is identified from (C.2).

The families of unconstrained youth who receive large amount of parental transfers also provide information about the EGS. As shown in Section 4.1, when the borrowing constraints do not bind, parental transfer decisions are primarily deriven by relative resources between parents and children, as well as the EGS, while the EIS has no effects. Since $\Pr(B^*|J, \boldsymbol{X}, \boldsymbol{Z})$ is identified, how parental transfers vary with youth ability as well as parental wealth identifies the EGS.

In contrast, the EIS is identified from borrowing-constrained youth who receive little parental tranfsers. To illustrate this more starkly, consider the schooling distribution of those who receive no parental support at all, i.e., $\Pr(J|B^* \in \{0\}, \boldsymbol{X})$, assuming $\{0\} \subset \mathcal{B}$. As shown by Lochner and Monge-Naranjo (2011), the EIS plays an important role in shaping the relationship between ability and schooling investment in the presence of borrowing constraints, for a given value of initial wealth (or parental transfer). For example, when there is a strong desire to smooth consumption over time, those with higher ability may invest less in schooling despite their higher returns.

More generally, at intermediate levels of parental transfers—where youth receive some support but remain borrowing-constrained—the consumption-smoothing parameters influence the relationship between educational attainment, parental transfers, and youth's ability. For example, a strong and positive ability-transfer gradient would suggest a large value of the EGS relative to the EIS. Moreover, for a given value of the EGS, the educational attainment–transfer gradient is informative about the EIS: with a small EIS, the intertemporal consumption distortion associated with further schooling is large, and therefore educational attainment depends strongly on parental transfers.

Finally, for given consumption-smoothing parameters and psychic returns to schooling parameters, the parental transfer function $\hat{b}_j(\delta, \boldsymbol{X})$ is a known function that is increasing in δ. Therefore, the conditional distribution of parental transfers $\Pr(B^*|\boldsymbol{X}, \boldsymbol{Z}, J)$ identifies the conditional distribution of parenal altruism $\Pr(\delta|\boldsymbol{Z})$. To see this, consider the following relationship:

$$\begin{aligned}
\Pr(B^* \in \mathcal{B}_{k^*}|\boldsymbol{X}, \boldsymbol{Z}, J) &= \sum_{m=1}^{M} \Pr(\delta = \delta_m, B^* \in \mathcal{B}_{k^*}|\boldsymbol{X}, \boldsymbol{Z}, J) \\
&= \sum_{m=1}^{M} \Pr(\delta = \delta_m|\boldsymbol{X}, \boldsymbol{Z}, J) \Pr(B^* \in \mathcal{B}_{k^*}|\boldsymbol{X}, \boldsymbol{Z}, J, \delta = \delta_m) \\
&= \sum_{m=1}^{M} \Pr(\delta = \delta_m|\boldsymbol{X}, \boldsymbol{Z}, J)\mathbb{I}_{\hat{b}_J(\delta_m, \boldsymbol{X}) \in \mathcal{B}_{k^*}}.
\end{aligned}$$

For a given $(\boldsymbol{X}, \boldsymbol{Z}, J)$, consider a $K \times M$ matrix with elements $\mathbb{I}_{\hat{b}_J(\delta_m, \boldsymbol{X}) \in \mathcal{B}_{k^*}}$. If the rank of this matrix is greater than or equal to M, $\{\Pr(\delta = \delta_m|\boldsymbol{X}, \boldsymbol{Z}, J)\}_{m=1}^{M}$ is identified from the equation above.

Next, notice that

$$\begin{aligned}\Pr(\delta = \delta_m|\boldsymbol{X}, \boldsymbol{Z}, J) &= \Pr(J|\delta = \delta_m, \boldsymbol{X}, \boldsymbol{Z}) \frac{\Pr(\delta = \delta_m|\boldsymbol{X}, \boldsymbol{Z})}{\Pr(J|\boldsymbol{X}, \boldsymbol{Z})} \\ &= \Pr(J|\delta = \delta_m, \boldsymbol{X}) \frac{\Pr(\delta = \delta_m|\boldsymbol{Z})}{\sum_{m=1}^{M} \Pr(\delta = \delta_m|\boldsymbol{X}, \boldsymbol{Z}) \Pr(J|\delta = \delta_m, \boldsymbol{X}, \boldsymbol{Z})} \\ &= \frac{\Pr(\delta = \delta_m|\boldsymbol{Z}) \Pr(J|\delta = \delta_m, \boldsymbol{X})}{\sum_{m'=1}^{M} \Pr(\delta = \delta_{m'}|\boldsymbol{Z}) \Pr(J|\delta = \delta_{m'}, \boldsymbol{X})}.\end{aligned}$$

From this, we get

$$\frac{\Pr(\delta = \delta_m|\boldsymbol{X}, \boldsymbol{Z}, J)}{\Pr(\delta = \delta_{m'}|\boldsymbol{X}, \boldsymbol{Z}, J)} = \frac{\Pr(\delta = \delta_m|\boldsymbol{Z}) \Pr(J|\delta = \delta_m, \boldsymbol{X})}{\Pr(\delta = \delta_{m'}|\boldsymbol{Z}) \Pr(J|\delta = \delta_{m'}, \boldsymbol{X})}.$$

Therefore, we can identify the ratios $\Pr(\delta = \delta_m|\boldsymbol{Z})/\Pr(\delta = \delta_{m'}|\boldsymbol{Z})$, and the restriction $\sum_{m=1}^{M} \Pr(\delta = \delta_m|\boldsymbol{Z}) = 1$ gives the levels $\{\Pr(\delta = \delta_m|\boldsymbol{Z})\}_{m=1}^{M}$.

C.4 Additional Details on Parameter Estimates

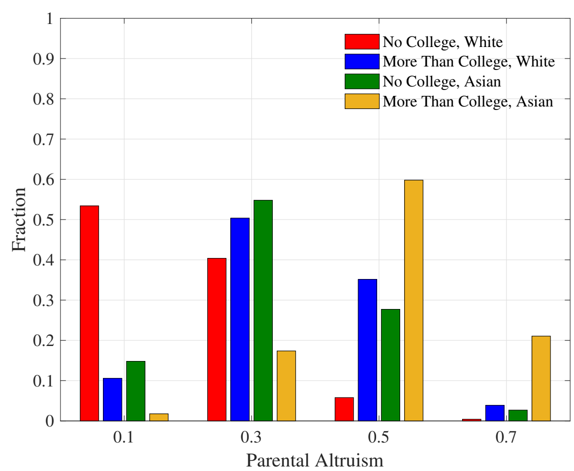


Figure C.1: Parental Altruism Distribution for Selected Groups ($\Pr(\delta|\boldsymbol{Z})$)

Figure C.1 presents the distribution of parental altruism for hypothetical one-parent, one-child families, highlighting the effects of parental education and race. White youth with parents who have not attended college exhibit the lowest levels of parental altruism, with a modal value of 0.1. Being Asian, or having parents with the highest level of education (i.e., more than college) raises the modal altruism to 0.3. The combination of being asian and having the most educated parents further increases the mode to 0.5.

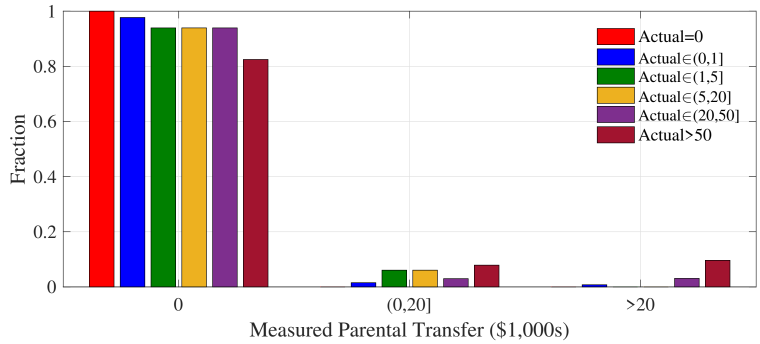


Figure C.2: Distribution of Secondary Parental Transfer Measures Conditional on Actual Transfer

D Details on Optimal Policy Simulations

D.1 Distortions in Quantitative Model

Distortions are defined analogously to Section 3.4. Conditional on $(\boldsymbol{b}, \boldsymbol{g}, a, h)$, the consumption distortion is given by:

$$L_c(\boldsymbol{b}, \boldsymbol{g}, a, h) := \sum_{j \in \mathcal{J}} p_j(\boldsymbol{b}, \boldsymbol{g}, a, h) \Big\{ \big[y_j(a) + b_j + g_j - k_j \big] - V_j(b_j + g_j; a) \Big\} \tag{D.1}$$

and the schooling distortion is

$$\begin{aligned} L_s(\boldsymbol{b}, \boldsymbol{g}, a, h) := & \int \max \left\{ y_0(a) + \varepsilon, \int \max_{j \in \mathcal{J}} \{ y_j(a) - k_j + \varepsilon_j \} dF_1(\boldsymbol{\varepsilon}_1 | a, h) \right\} dF(\varepsilon) \\ & - \int \mathbb{I}_{\varepsilon \geq V_1(\boldsymbol{b}_1, \boldsymbol{g}, a, h) - V_0(b_0; a)} \big[y_0(a) + \varepsilon \big] dF(\varepsilon | a, h) \\ & - p(\boldsymbol{b}, \boldsymbol{g}, a, h) \sum_{j \in \mathcal{J}} \int \mathbb{I}_{j \in \operatorname{argmax}_{j' \in \mathcal{J}} \{ V_{j'}(b_{j'} + g_{j'}; a) + \varepsilon_{j'} \}} \big[y_j(a) - k_j + \varepsilon_j \big] dF_1(\boldsymbol{\varepsilon}_1 | a, h), \end{aligned}$$

where $\boldsymbol{\varepsilon}_1 := (\varepsilon_j)_{j \in \mathcal{J}}$ and $F_1(\boldsymbol{\varepsilon}_1 | a, h)$ is the cumulative distribution function of $\boldsymbol{\varepsilon}_1$ conditional on (a, h).

Adding these two gives the total distortion

$$\begin{aligned} L(\boldsymbol{b}, \boldsymbol{g}, a, h) = & \int \max \left\{ y_0(a) + \varepsilon, \int \max_{j \in \mathcal{J}} \{ y_j(a) - k_j + \varepsilon_j \} dF_1(\boldsymbol{\varepsilon}_1 | a, h) \right\} dF(\varepsilon) \\ & + p_0(\boldsymbol{b}, \boldsymbol{g}, a, h) b_0 + \sum_{j \in \mathcal{J}} p_j(\boldsymbol{b}, \boldsymbol{g}, a, h)(b_j + g_j) - V(\boldsymbol{b}, \boldsymbol{g}, a, h). \end{aligned}$$

Therefore, for a given level of $p_j(\boldsymbol{b}, \boldsymbol{g}, a, h) g_j$, minimizing $L(\boldsymbol{b}, \boldsymbol{g}, a, h)$ is equivalent to maximizing

$$w - \sum_{j\in\overline{\mathcal{J}}} p_j(\boldsymbol{b}, \boldsymbol{g}_h, a, h) b_j + V(\boldsymbol{b}, \boldsymbol{g}_h, a, h) = \sum_{t=1}^{T_p} R^{-t} c_p + \sum_{t=1}^{T_k} R^{-t} c_k.$$

where c_p and c_k are defined in (13) and (14).

D.2 Numerically Solving for Optimal Policy

I solve this constrained optimization problem numerically using an iterated two-step algorithm. Let $\lambda \geq 0$ be the Lagrange multiplier on the social budget constraint (21). In the inner loop, for a given value of λ, I solve for $(\boldsymbol{g}_h)_{h\in\mathcal{H}}$ that minimizes the Lagrangian,

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{m=1}^{M} \Pr(\delta = \delta_m | \boldsymbol{Z}_i) \left[L\big(\hat{\boldsymbol{b}}(\delta_m, \boldsymbol{X}_i), \boldsymbol{g}_{h_i}, a_i, h_i\big) + \lambda \sum_{j\in\mathcal{J}} p_j\big(\hat{\boldsymbol{b}}(\delta_m, \boldsymbol{X}_i), \boldsymbol{g}_{h_i}, a_i, h_i\big) g_{j,h_i} \right],$$

subject to the constraints $g_{j,h} \geq 0$ for all $(j,h) \in \mathcal{J}\times\mathcal{H}$ and (18). In the outer loop, I search for λ such that the social budget constraint (21) holds with equality. As noted earlier, this constraint may be slack, in which case $\lambda = 0$. Therefore, I first check whether the constraint is satisfied at $\lambda = 0$ before imposing equality.

E Additional Details on Section 6

E.1 Analytical Details for Section 6.1

The efficiency loss from suboptimal schooling choices that accounts for the social return is

$$\tilde{L}_s(b,g) := \int \max\{\tilde{y}_0 + \varepsilon, \tilde{y}_1 - k\} - \left\{\mathbb{I}_{\varepsilon > V_1(b+g) - V_0(b)}(\tilde{y}_0 + \varepsilon) + \mathbb{I}_{\varepsilon \leq V_1(b+g) - V_0(b)}(\tilde{y}_1 - k)\right\} dF(\varepsilon),$$

Then, the total efficiency loss in this case can be written as follows:

$$\tilde{L}(b,g) = \int \max\{\tilde{y}_0 + \varepsilon, \tilde{y}_1 - k\} dF(\varepsilon) + b - \big[1 - p(b,g)\big]\tau_0 + p(b,g)(g - \tau_1) - V(b,g).$$

E.2 Planning Problem with Externality

Consumption distortion remains as defined in (D.1). As in Section 6.1, schooling distortions are measured based on before-tax earnings:

$$
\begin{aligned}
\tilde{L}_s(\boldsymbol{b}, \boldsymbol{g}, a, h) := & \int \max \left\{ \tilde{y}_0(a) + \varepsilon, \int \max_{j \in \mathcal{J}} \left\{ \tilde{y}_j(a) - k_j + \varepsilon_j \right\} dF_1(\boldsymbol{\varepsilon}_1 | a, h) \right\} dF(\varepsilon) \\
& - \int \mathbb{I}_{\varepsilon \geq V_1(\boldsymbol{b}_1, \boldsymbol{g}, a, h) - V_0(b_0; a)} \left[\tilde{y}_0(a) + \varepsilon \right] dF(\varepsilon | a, h) \\
& - p(\boldsymbol{b}, \boldsymbol{g}, a, h) \sum_{j \in \mathcal{J}} \int \mathbb{I}_{j \in \operatorname{argmax}_{j' \in \mathcal{J}} \{ V_{j'}(b_{j'} + g_{j'}; a) + \varepsilon_{j'} \}} \left[\tilde{y}_j(a) - k_j + \varepsilon_j \right] dF_1(\boldsymbol{\varepsilon}_1 | a, h),
\end{aligned}
$$

where $\tilde{y}_j(a)$ denotes the lifetime before-tax earnings associated with schooling option j:

$$
\tilde{y}_j(a) := \sum_{x=1}^{65-17-T_j} R^{-(x+T_j)} \tilde{y}_j(a, x),
$$

Total inefficiency is then defined as $\tilde{L}(\boldsymbol{b}, \boldsymbol{g}, a, h) := \tilde{L}_s(\boldsymbol{b}, \boldsymbol{g}, a, h) + L_c(\boldsymbol{b}, \boldsymbol{g}, a, h)$.

Let $\tau_j(a)$ denote the lifetime tax payment associated with schooling option j and ability level a:

$$
\tau_j(a) := \sum_{x=1}^{65-17-T_j} R^{-(x+T_j)} \mathcal{T}\big(\tilde{y}_j(a, x)\big).
$$

The planner's problem is:

$$
\begin{aligned}
& \min_{(\boldsymbol{g}_h)_{h \in \mathcal{H}}} \frac{1}{N} \sum_{i=1}^{N} \sum_{m=1}^{M} \Pr(\delta = \delta_m | \boldsymbol{Z}_i) \tilde{L}\big(\hat{\boldsymbol{b}}(\delta_m, \boldsymbol{X}_i), \boldsymbol{g}_{h_i}, a_i, h_i\big) \\
& \text{subject to } \frac{1}{N} \sum_{i=1}^{N} \sum_{m=1}^{M} \Pr(\delta = \delta_m | \boldsymbol{Z}_i) \sum_{j \in \mathcal{J}} p_j\big(\hat{\boldsymbol{b}}(\delta_m, \boldsymbol{X}_i), \boldsymbol{g}_{h_i}, a_i, h_i\big) g_{j, h_i} + E \leq \\
& \qquad \frac{1}{N} \sum_{i=1}^{N} \sum_{m=1}^{M} \Pr(\delta = \delta_m | \boldsymbol{Z}_i) \sum_{j \in \overline{\mathcal{J}}} p_j\big(\hat{\boldsymbol{b}}(\delta_m, \boldsymbol{X}_i), \boldsymbol{g}_{h_i}, a_i, h_i\big) \tau_j(a_i),
\end{aligned}
$$

together with the constraints $g_{j,h} \geq 0$ for all $(j, h) \in \mathcal{J} \times \mathcal{H}$ and constraint (18), where $\hat{\boldsymbol{b}}(\delta, \boldsymbol{X})$ solves the parent's problem for a family with $(\delta, \boldsymbol{X})$. The amount E, calculated under the current policy, is held fixed when solving for the optimal policy.