The Welfare Cost of Under-utilization of Need-based College Financial Aid*

Emily G. Moschini† Gajendran Raveendranathan‡

May 30, 2024

Abstract

In the United States, high school graduates from poor families have lower college enrollment rates compared to their higher-income peers with similar high school grades. This is true even for those with grades in the highest grade tercile. At the same time, a sizable federal financial aid system seeks to help low-income young people finance college education. Contributing to this enrollment pattern may be that young people misunderstand their eligibility for need-based postsecondary financial aid, or that applying for aid is prohibitively costly. In this paper, we use nationally representative panel data from the US Department of Education to quantify the extent of mistaken beliefs about financial aid eligibility, and find that 10 percent of high school graduates who do not enroll in any postsecondary education are misinformed about their eligibility for federal grants. Using the same data, we find that 6 percent of nonenrollees would have been eligible for federal grant aid, but cite application fixed costs as a reason for not applying for this aid. We embed financial aid frictions into a life cycle model that includes federal grants, subsidized and unsubsidized federal student loans, endogenous parental transfers, and earnings from part-time work. After calibrating our model to match the estimated share of high school graduates who do not enroll in college and who have mistaken beliefs about financial aid eligibility, we shut off mistaken beliefs in the model to quantify their welfare impact. We find that their presence generates welfare costs for higher-skill young adults in partial equilibrium, while general equilibrium adjustments lead to losses for all skill levels.

JEL classification numbers: I22, I26, E7, G28, G5
Keywords: Subjective beliefs, Information intervention, Inter-generational transfers.

*Raveendranathan gratefully acknowledges financial support from the Social Sciences and Humanities Research Council (SSHRC) via Insight Grant 435-2024-0307.

†William & Mary. E-mail: egmoschini@wm.edu.
‡McMaster University. E-mail: raveeg1@mcmaster.ca.

1 Introduction

In the United States, low-income high school graduates are less likely to enroll in postsecondary education than their richer peers with similar high school GPAs. This is true even for those with grades in the top tercile. At the same time, a large federal financial aid system exists, a significant portion of which is explicitly means-tested and seeks to encourage postsecondary enrollment by those from poor families. Frictions in the process for accesing financial aid may contribute to this pattern of enrollment in the face of available financial aid. For example, young people may be mistaken about their own eligibility for need-based financial aid and thus not apply for it, or may face prohibitive fixed costs to accessing this aid (e.g., due to complicated paper work). In this paper, we focus on the decision to enroll in a bachelor's degree program in an environment with a financial aid system similar to that of the United States in the mid-2010s. We provide new empirical evidence on the extent of mistaken beliefs about eligibility for need-based federal financial aid (in particular Pell grants). We also provide new evidence on the extent of prohibitive fixed costs when applying for this aid. To quantify the cost of mistaken beliefs about eligibility, we embed them in a structural model that includes an approximation of the expected family contribution (EFC) formula, a key ingredient for determining financial aid eligibility during this period. We find that, in partial equilibrium, the costs of mistaken beliefs are borne by those with high skill, while in general equilibrium all skill levels experience welfare losses. The highest welfare costs are large, and accrue to young people with high skill with poor parents and mistaken beliefs.

Our main empirical results are established using the High School Longitudinal Study of 2009, a panel survey from the United States Department of Education. We use this data to estimate eligibility for need-based financial aid by approximating the complex system that determined eligibility in the mid-2010s; in particular, we compute estimates of the EFC by imputing it using this public data, which then allows us to estimate financial aid eligibility for those who never receive aid. Using our estimates of eligibility and directly reported reasons given by survey respondents for not applying for financial aid, we find that 10 percent of high school graduates who do not enroll in any postsecondary education after high school did not apply for financial aid because of mistaken beliefs about their eligibility. In addition, 6 percent of non-enrollees did not apply due to fixed costs but would have been eligible for Pell grant aid if they had done so.

In order to quantify the economic impact of mistaken beliefs about need-based financial aid eligibility, we embed beliefs about eligibility which may be incorrect into a structural model of college enrollment choice. This model also features a financial aid policy environment similar to that of the United States in the mid-2010s. In particular, our model approximates the way that family attributes map into eligibility for need-based financial aid via the EFC; sources of college financial aid include Pell grants, subsidized federal loans, and unsubsidized federal loans. We also include endogenous family transfers and earnings from part-time work as other sources of financing for college expenses.

After calibrating our framework to match the estimated extent of mistaken beliefs in the population of high school graduates, we validate the model with five exercises. First, we show that the model aligns well with the data when we compare the model's enrollment responses to a tuition subsidy with estimates from Deming and Dynarski (2009). Next, we compare the extent of "under-enrollment" in college in the model with our estimates from the data, where under-enrollment means that young people do not enroll but would have done so with correct beliefs (all else being equal). We find that the model is conservative with respect to the data in its extent of under-enrollment. In our third exercise, we compute the college wage premium by skill tercile in the model and compare it with our estimates of its empirical counterpart; the model moments are a close match with the data. In our fourth exercise we compare enrollment rates by skill and parental income terciles in the model baseline with empirical estimates. The model captures the qualitative pattern in the data: enrollment is increasing in skill and parental income. However, the model understates enrollment for those outside of the top parental income and skill terciles. In our final exercise, we compare model need-based financial aid uptake with the data, along both the extensive and intensive margin. The model underestimates uptake rates (intensive margin) but is close to the data in terms of the amount of aid conditional on receipt (extensive margin).

To quantify the welfare cost of misinformation, we perform an "information intervention," in which the beliefs of young adults are corrected before the college enrollment decision; we focus on 18-year-olds because the target population of information interventions are high school graduates. To compute welfare changes, we calculate consumption-equivalent variation for 18-year-olds, taking consumer choices as given. We then examine welfare changes over the transition to the new steady-state, and find that eliminating mistaken beliefs generates welfare gains for 18-year-olds with high skill in partial equilibrium. This group benefits because their enrollment decisions change once beliefs are corrected while other skill levels are less responsive. In general equilibrium, as the misallocation caused by mistaken beliefs is corrected, the return to labor supply increases for adults without a college degree. This adjustment leads to welfare gains even for those with lower skill levels, whose enrollment choices remain largely unchanged compared to the baseline. Welfare gains for the highest skill level are dampened but nevertheless remain after general equilibrium adjustments are taken into account. The gains from eliminating mistaken beliefs are highest among young people with high skill, poor parents, and mistaken beliefs. Overall, because eliminating mistaken beliefs leads to welfare gains, these results indicate that the presence of mistaken beliefs about financial aid eligibility generates welfare costs in the baseline economy.

Our results offer an empirically-disciplined analysis of the potential macroeconomic consequences of providing factual information to young people when they make education choices early in adulthood. Our main experiment can be viewed as a scaled-up version of interventions examined by Hoxby and Turner (2015) and Dynarski, Libassi, Michelmore, and Owen (2021), although those studies focused on beliefs about slightly different aspects of postsecondary enrollment and financing. We find that the highest gains from the interformation intervention accrue to a group of young people whose attributes broadly align with the target population for these previous small-scale interventions: academically gifted, relatively poor, and initially lacking information about the costs and benefits of their postsecondary education options.

The paper proceeds as follows. Section 2 overviews the financial aid system in the United States. Section 3 explains our main empirical results, and Section 4 outlines our structural model which is parameterized in Section 5. We explain and analyze results from our main experiment in Section 6. Section 7 concludes.

2 Postsecondary Financial Aid in the United States

Federal Student Aid in the United States is a system of public financial support for those pursuing education after high school. This section briefly outlines the institutional mechanisms that allocated financial student aid in the period covered by our data, with the goal of lending clarity to our empirical analysis in Section 3 and our model framework in Section 4.

The extent of financial aid that an applicant is eligible for is largely based on the inputs into the Free Application for Federal Student Aid (FAFSA) form, which students complete while submitting applications to postsecondary institutions and before each academic year of enrollment thereafter (Collins, 2021). Submitting the FAFSA form is a necessary condition for financial aid receipt. Previous research has highlighted the complexity of the FAFSA form and, more broadly, the US financial aid system (Dynarski and Scott-Clayton, 2006).

The FAFSA collects information on the financial resources of an applicant's family. Financial resources are measured by the FAFSA using income in the previous year and some information on savings; the FAFSA then makes allowances for federal tax liability, parent employment status, and number of other dependents currently enrolled in college. The output of the FAFSA is the expected family contribution (EFC), which is the amount of money that the family is expected to contribute towards college expenses.[1] The difference between the COA and the EFC is the "unmet need" of

[1]Parental financial resources are required for dependent filers; alternatively, students may file as an independent. Most first-year students who apply for financial aid do so as a dependent of their parents. In this paper, we focus on analyzing dependent filers. For dependents, technically the FAFSA also assigns an expected contribution of the

the student, which may be partially or fully met by financial aid. When students who submitted a FAFSA hear back from postsecondary institutions about the outcome of their application for admittance, they are also offered a financial aid package drawn up by employees of the institution using information on the FAFSA and knowledge of that institution's cost of attendance as well as knowledge about the student.[2]

Federal financial aid takes two main forms: grants and loans.[3] The main type of federal grant is the Pell grant. Federal student loans are mostly Stafford loans, which are direct loans to students from the federal government. Stafford loans are either unsubsidized or subsidized: for the former type, interest accrues on the loan balance while the recipient is enrolled, while for the latter type it does not. Both Pell grants and Stafford loans require a FAFSA form submission for receipt to be possible; subsidized Stafford loans and Pell grants are additionally means-tested in their eligibility, while unsubsudized Stafford are available to any postsecondary student regardless of need (Dortch, 2023). In practice, therefore, the student loan component of the US federal postsecondary financial aid system is a combination of means-tested and universal eligibility. In the economic literature on this topic, the term "financial aid" can be used to refer to the entire financial aid system, which is partially means-tested, or to the means-tested components of the financial aid system (e.g., subsidized Stafford loans and Pell grants). When possible, we are specific in this paper about which component of financial aid we are referring to in our analysis (e.g., means-tested or not).

The magnitude of Pell grants offerred to a student is determined by the difference between an upper bound and their EFC, with the upper bound set as the minimum of a statutory maximum Pell grant limit and the COA assigned to the student. The annual borrowing limit for federal Stafford loans is determined using a combination of the COA, the EFC, Pell grants, and statutory annual limits on federal loans. Statutory annual borrowing limits for Stafford loans are increasing in one's academic year; there is also a cumulative lifetime borrowing limit. Subsidized Stafford loans have annual and cumulative borrowing limits below the limits for total federal loans.

Financial aid packages may include not only federal aid, but state aid and campus-based aid such as

student, and which we will abstract from in this paper (it is usually much lower than the expected parent contribution that we focus on).

[2]Only postsecondary institutions who meet certain criteria are permitted to administer federal financial aid. These institutions may be less than 2-year programs, such as those issuing certificates (e.g., hairdresser), 2-year programs (associate's degrees), or 4-year programs (bachelor's degrees). In this paper, we use the term "2-year program" interchangeably with "associate's degree program," and "4-year program" interchangeably with "bachelor's degree program". In order to remain eligible for disbursement of federal financial aid, the institution must meet certain thresholds regarding student outcomes after graduation.

[3]Campus-based financial aid programs during the period we study included federal work-study, Perkins loans, and the Federal Educational Opportunity Grant program. These programs are relatively small compared to Pell grants and Stafford loans, and eligibility/allocation is also determined according to formulae specific to the institution rather than broad rules set by statute (Congressional Research Service of the United States [author name redacted], 2018). In our analysis we abstract from these programs and focus on Pell grants and Stafford loans (subsidized or unsubsidized).

institutional grants. The components of a financial aid package are a-la-carte: potential enrollees may reject the loan component of the package, for example, but request the grant component. Given their financing options at the institutions that admit them, applicants then decide which college to enroll in. In each subsequent year of enrollment, students must re-submit the FAFSA with updated information, and their financial aid package is re-determined at the institution where they have enrolled before the start of the next academic year.

The composition of financial aid packages that postsecondary institutions offer students is determined by Financial Aid Administrators (FAAs) who work at those educational institutions and who draw up packages according to guidelines set out in the Higher Education Act (NASFAA Monograph, 8th Edition, 2023; Office of Postsecondary Education, Federal Student Aid, 2022). FAAs have considerable discretion in designing the composition of student aid packages: for example, they can adjust the estimated cost of attendance for the FAFSA filer and choose whether to emphasize various types of aid in the package that they offer to each applicant. The sources of aid that a given FAA can draw on varies across institutions although all accredited institutions may offer Pell grants and Stafford loans. In addition, FAAs are not required to make sure that the unmet need of applicants is covered by the aid package they are offered.

Because of these aspects of the financial aid system in the United States, there is considerable idiosyncratic variation in many aspects of financial aid even for those with the same EFC and similar student attributes. This variation is introduced by variation across institutions in cost of attendance, and also by variation within the same institution in the way the FAA processes each applicant. Nevertheless, several components of financial aid are dictated by federal policy in a predictable way. For example, when designing financial aid packages, the FAA is required to prioritize Pell grants; once the amount of Pell grants one is eligible for is exhausted, the FAA turns to other sources such as subsidized Stafford loans or campus-based aid such as work-study or institutional grants, or unsubsidized Stafford loans.

In the next section, we describe our main source of data and explain how we impute eligibility for Pell grants in a sample of recent high school graduates. We are especially interested in eligibility for Pell grants because, among the sources of financial aid, it is least affected by the idiosyncratic variation introduced by the role of FAAs in the financial aid allocation process. After validating the imputation against a subsample who submitted a FAFSA and enrolled in a bachelor's program, we quantify the extent of confusion about eligibility for means-tested federal aid in the 2013 cohort of US high school graduates by comparing imputed financial aid eligibility with self-reported student beliefs. We also document that a proportion of high school graduates report not filing a FAFSA due to the difficulty of the form, reflecting an additional fixed cost friction.

3 Data

Our main data source is the public version of the High School Longitudinal Study of 2009 (HSLS:09), a nationally representative panel survey of 9th graders in the United States collected by the National Center for Education Statistics (NCES), a subsidiary of the Department of Education (National Center for Education Statistics, U.S. Department of Education, 2020). There are four main waves of data collection in the HSLS:09: the base year (2009), first follow-up (2012), 2013 update (2013), and second follow up (2016), supplemented by Student Records and Postsecondary Transcripts collected from institutional sources around the 2017 academic year for postsecondary enrollees. Additionally, financial aid records are pulled from the National Student Loan Database for the sample with a student record collection. In the base year and first follow-up, the sample member and their parent are both interviewed, while the 2013 Upate is an abbreviated survey questionnaire completed by the sample member or their parent (in the latter case, survey questions are reframed to be about their child). Only the sample member is interviewed in the second follow-up. A useful feature of the HSLS:09 is that it collects information on all sample members after high school regardless of postsecondary education outcome.

Summary statistics We begin by restricting attention to high school graduates for whom we have a consistent measure of postsecondary enrollment outcomes by 2016 and for whom we observe family income, honors-weighted high school GPA, and whether or not the respondent submitted a FAFSA in the 2013 Update.[4] We compare this "cleaned" sample with the original "raw" data in Table 1. Appendix A.1.1 provides further details of our sample selection procedure as well as an extended description of how we assign observations to postsecondary outcomes by capitalizing on the various sources of information available in the HSLS:09.

Comparing the raw and cleaned samples in Table 1 indicates that the cleaned sample has fewer observations, which partially affects sample composition compared to the raw sample. Panel A reports student attributes: the cleaned sample exhibit higher high school GPA and fewer females. Family attributes are reported in Panel B: the cleaned sample exhibit a higher average family income, a lower poverty rate, a greater share of two-parent families, a slightly lower household size, and a slightly higher number of siblings already enrolled in a postsecondary program when the sample member is about to finish high school. Panel C contains rates of postsecondary enrollment outcomes, which are not directly available in the raw data: in the cleaned sample, 45 percent of

[4]Honors-weighted GPA takes into account whether a given course is honors or college level, and then adjusts how the grade received in that course contributes to overall GPA accordingly. For a given letter grade, this adjustment raises the total GPA of the individual if the difficulty of the courses taken is higher. See HSLS:09 2013 Update: Student file codebook and HSLS:09 2013 Update User Manual Appendix for the specific formulas used to compute honors-weighted high school GPA in the HSLS:09.

high school graduates enroll in a bachelor's degree, 21 percent in a 2-year program, and 8 percent in a certificate program or undergraduate courses (without pursuing a specific degree). About 25 percent of this cohort do not enroll in any postesecondary education. Table A1 of Appendix A.1.1 compares enrollment rates in 2- and 4-year programs that we compute in the HSLS:09 cleaned sample with national-level statistics for the 2013-2014 academic year reported by the NCES, and shows that they are quite consistent. This lends support to the accuracy of our postsecondary outcome assignment procedure.[5]

In Table A2 of Appendix A.1.1 we report descriptive statistics by each postsecondary education outcome such as the average high school GPA of enrollees and tuition and fees, as well as later transitions into different postsecondary education outcomes by 2016, broken down by 2013 postsecondary enrollment status. Table A3 reports BA enrollment by respondent terciles of respondent skill and parent income, demonstrating that enrollment is increasing in family income for each skill tercile. This pattern also holds true when we include both 2- and 4-year programs in our enrollment rates, as reported in Table A4. Lower enrollment rates for students from poor families may be related to frictions in accessing aid through the means-tested federal financial aid system. We are especially interested in how enrollment in bachelor's degree programs interacts with financial aid, because BAs have higher wage premia than other programs, on average, and they are also more expensive to complete than other types of post-secondary education. Distortions in enrollment and financing for BAs is therefore more likely to have large implications than other post-secondary outcome margins. One factor driving differences in BA enrollment across family income levels may be mistaken beliefs about eligibility for federal financial aid or fixed costs of completing a FAFSA.

FAFSA nonfilers and counterfactual Pell grant aid In order to empirically discipline why potential postsecondary students do not apply for financial aid, we turn to the 2013 Update questionnaire. Regardless of postsecondary enrollment outcome, the 2013 Update asks the sample member (or their parent) whether or not they completed the FAFSA for themselves (or on behalf of the sample member). If a FAFSA was not completed, the questionnaire asks "Why?" and provides a list of 7 possible reasons for not submitting the form. The survey then asks the respondent to agree or disagree with each possible reason; respondents may agree with more than one reason. This list is: not wanting to go into debt, being able to afford college without financial aid, believing themselves ineligible or unqualified, not knowing how to complete the FAFSA, because the forms were too much work, not knowing that they could, and not planning to continue their education. Appendix A.1.4 provides specific phrasing and additional details on these questions.

[5] Table A5 in Appendix A.1.1 compares the cleaned sample with a conditional sample where those enrolled in less than a 4-year postsecondary education program are not included.

Table 1: Summary statistics

		Sample	
Category	**Variable**	**Raw**	**Cleaned**
Panel A: Student attributes	High school GPA	2.816	3.040
	Female	0.506	0.502
Panel B: Family attributes	Family income	78,496	88,305
	Below 185 % of poverty line	0.394	0.329
	Two-parent family	0.873	0.875
	Household size	4.269	4.248
	Number siblings in PSE	0.360	0.422
Panel C: Enrollment outcome in fall 2013	Did not enroll	-	0.253
	Less than 2-year	-	0.079
	Associate's (AA, 2-year)	-	0.214
	Bachelor's (BA, 4-year)	-	0.453
	Observations	13,283	7,045

Notes: Table 1 reports means of variables for the raw and cleaned samples from the HSLS:09. Weights: Second Follow-up student longitudinal weights. Source: HSLS:09.

We group the 7 possible reasons into three categories: first, reasons related to beliefs about financial aid eligibility; second, reasons related to the fixed cost of filing a FAFSA; and, third, other reasons not related to frictions. Specifically, we group the responses "believing themselves ineligible or unqualified' or "because did not know could' or "because did not want to go into debt"' as responses that indicate that the respondent's understanding of their eligibility for financial aid was a reason that they did not submit a FAFSA. Not wanting to go into debt is included in this category because financial aid is not synonymous with student debt due to the Pell grant program. We group "did not know how" and "because the forms were too much work" as reflecting fixed costs of completing a FAFSA. Finally, we consider "can afford college without financial aid" and "not planning to continue education" as other reasons not related to frictions. In each case, we create an indicator equal to 1 if the respondent indicated that a FAFSA non-filing reason applied to them, and set to zero otherwise.

Next, we simulate eligibility for Pell grants for FAFSA nonfilers. This exercise approximates the output of a FAFSA form (that is, the EFC) for the cleaned sample of Table 1, drawing on descriptions of the 2013-2014 FAFSA form available in the EFC formula guide for 2013-2014 prepared by the Federal Student Aid Office, U.S. Department of Education (2014). Along with imputed COA, we use the estimated EFC to impute the level of Pell Grants that the respond would be eligible for were they to enroll in a BA program, using descriptions from a Congressional Research Service report (Dortch, 2023).[6] Details of our imputation procedure are provided in

[6]Financial aid partly determined by cost of attendance; we construct counterfactual financial aid using bachelor's degree cost of attendance to be consistent with the two postsecondary outcomes of our model environment: not enrolling, or enrolling in a bachelor's degree program.

Table A9 of Appendix A.1.3. In Table A10 of that appendix, we validate the accurate of our imputation by comparing imputed Pell grants to realized Pell grants among FAFSA filers who enroll in a bachelor's degree in the fall of 2013. Our imputation provides a good fit with realized counterparts for this subsample: for observations where imputed and realized Pell grants are both observed, the median difference between realized and imputed Pell grants is 0; for observations where the imputed Pell grant is required to be positive, the median difference is -$50.

Panel A of Table 2 first reports the frequency of not filing a FAFSA by postsecondary outcome among high school graduates for the cleaned sample. The FAFSA nonfiling frequency is highest among those who do not enroll in any postsecondary education. The second variable is the share of high school graduates who do not file a FAFSA due to reasons related to beliefs about financial aid eligibility. Almost 20 percent of high school graduates who did not enroll in any postsecondary education indicate that this is a reason that they did not file a FAFSA; the number is decreasing but nonzero as one moves across columns towards the highest postsecondary enrollment outcome (BA program). The third variable is the share of high school graduates who are FAFSA nonfilers and say that they did not apply for aid due to fixed cost reasons. About 11 percent of non-enrollees meet this criteria. Next, incorporating information from our counterfactual Pell grant simulation, we report the share of high school graduates who did not file a FAFSA due to believing themselves ineligible, but who would have been eligible for a positive amount of Pell grants according to our imputation. This share is about 10 percent of those who do not enroll in any postsecondary education, and almost zero among BA students. Those who did not file due to fixed costs but were eligible for a positive Pell grant amount to 6 percent of non-enrollees and almost zero percent of BA students.

In Panel B, we report the conditional mean of imputed Pell grant eligibility, for the sample who did not file a FAFSA due to believing themselves ineligibility for financial aid and for those who did not apply for aid due to fixed costs of filing a FAFSA. For both groups, the magnitude of the money "left on the table" by FAFSA nonfilers is sizable relative to the cost of attendance (which we report in Table A2). This panel flags cells where the conditional sample is quite small.

In Table A11 of Appendix A.1.4, we document that believing oneself ineligible increases in frequency as family income increases. If this response to reasons for not filing a FAFSA was only noise, it would be less likely to be qualitatively consistent with the means-tested nature of the financial aid system in this way. The breakdown also indicates that the frequency of nonfiling due to *mistaken* beliefs about financial aid eligibility is highest for those from poorer families, which is expected given the means-tested nature of Pell grant eligibility.

Note that focusing on eligibility for Pell grants represents a lower bound of mistaken beliefs about means-tested financial aid eligibility; if we impute eligibility for subsidized Stafford loans as equal to the remaining unmet need after Pell grants are allocated (imposing additional limits on overall borrowing from federal loans), then this would further raise the frequency of mistaken beliefs, albeit with the caveat that the final role of subsidized loans in each aid package is more subject to idiosyncratic variation than the role of Pell grants is. In addition, Panel A of Table A12 in Appendix A.1.4 reports the share of high school graduates who are FAFSA nonfilers and who satisfy a more restrictive definition of believing themselves ineligible for financial aid. The frequency of mistaken beliefs in among high school graduates remains sizable, at about 7 percent. In Panel B of the same table, we report the share of high school graduates who say they do not file a FAFSA for either reasons related to beliefs about eligibility or fixed costs; 12 percent of the population does not apply for financial aid for at least one of these reasons but would be eligible for a Pell grant if they had. Panels C and D examine frequencies of those who say they did not apply for financial aid because of only believing themselves ineligible or only because of fixed costs. These frequencies are lowered relative to when we allow respondents to also include other reasons, but they are statistically different from zero.

Compositional differences between 2013 non-enrollee FAFSA nonfilers who are misinformed and those who are not are compared in Table A13; one takeaway of that table is that the misinformed are less likely to later enroll in post-secondary education. Table A14 performs the same comparison for those who report being affected by the fixed cost of filing and reaches similar conclusions. Additionally, Table A15 reports the share of FAFSA nonfilers who later are observed having applied for financial aid (by 2016). This frequency is less than 1 percent.

Attributes of our model environment In this paper, we focus on Pell grants and Stafford loans (subsidized or unsubsidized) as sources of federal financial aid for financing bachelor's degrees. The focus of our model framework on bachelor's degree programs, and the inclusion of all three types of federal financial aid, is partially motivated by the results of a financial aid accounting exercise presented in Table A17 of Appendix A.1.6. In particular, we report the distribution of federal financial aid in the aggregate (i.e., aggregating for the HSLS:09 2013 cohort of postsecondary enrollees), both across and within postsecondary education types. Our findings from this financial aid accounting exercise indicate that the majority of each type of aid is spent on bachelor's degree programs and that financial aid spent on bachelor's degree programs is spent roughly evenly on Pell grants, subsidized Stafford loans, and unsubsidized Stafford loans. We have also shown here that a sizable minority of high school graduates who do not enroll in post-secondary programs after high school are mistaken about their eligibility for Pell grants (the most predictable component of financial aid packages, given the financial aid system outlined in Section 2). A sizable share are also facing fixed costs that they view as affecting their financial aid application decision. Moti-

Table 2: FAFSA nonfilers: reasons, mistaken beliefs, and counterfactual Pell grant aid

		Enrollment status in fall of 2013			
Description	**Variable**	**Did not enroll**	**l.t. 2-year**	**AA (2-year)**	**BA (4-year)**
Panel A: High school graduates	FAFSA nonfilers	0.6240	0.3138	0.2490	0.1244
		(0.01297)	(0.02129)	(0.01221)	(0.00527)
	...Believe ineligible	0.1953	0.1582	0.0888	0.0773
		(0.01062)	(0.01675)	(0.00803)	(0.00427)
	...Fixed costs	0.1130	0.0803	0.0346	0.0174
		(0.00848)	(0.01247)	(0.00516)	(0.00209)
	...Believe ineligible and $\widehat{\text{Pell}} > 0$	0.1014	0.0544	0.0300	0.0049
		(0.00809)	(0.01041)	(0.00482)	(0.00112)
	...Fixed cost and $\widehat{\text{Pell}} > 0$	.06235	.03487	.01456	.00158
		(6.5e-03)	(8.4e-03)	(3.4e-03)	(6.3e-04)
Panel B: $\widehat{\textbf{Pell}} \mid \widehat{\textbf{Pell}} > 0$ and nonfiling reason	Believe ineligible	5,076	5,645 †	4,875 †	3,816 †
		(131)	(0)	(246)	(407)
	Fixed costs	4,903	5,645 †	4,056 †	3,716 †
		(182)	(0)	(380)	(699)
	Observations	1,395	476	1,256	3,918

Notes: For high school graduates, Panel A of Table 2 reports the share of FAFSA nonfilers, nonfilers because they believed themselves ineligible, nonfilers becaues of fixed costs of filing, and nonfilers giving each reason who would have counterfactually received a positive Pell grant according to our imputation. Panel B reports conditional means for imputed Pell grant amounts by nonfiling reason. Sample: cleaned sample from Table 1. Standard errors are in parentheses; dollar-value variables are in 2012 US dollars; observation counts refers to the number of high school graduates in a given postsecondary outcome, not the number of observations in the conditional sample in Panel B; † indicates cells where sample size is less than 50 observations. Weights: Second Follow-up student longitudinal weights. Source: HSLS:09.

vated by these findings, we include both information and fixed cost frictions in the financial aid application process in our model environment.

4 Model

We extend the framework of Moschini, Raveendranathan, and Xu (2023) by incorporating potentially mistaken subjective beliefs about eligibility for need-based college financial aid. As a true counterpart to beliefs about financial eligibility, we also include a model equivalent for the EFC formula that determines true eligibility, as well as Pell grants and subsidized federal student loans as distinct sources of college financial aid. This paper abstracts from mistaken beliefs about graduation likelihood, the main focus of Moschini, Raveendranathan, and Xu (2023).

Time is discrete and runs forever; each period lasts one year. The economy contains a government, a final goods firm, and heterogeneous consumers. This section overviews the agent problems and presents a subset of consumer value functions; the remaining value functions are provided in Appendix B.1. Equilibrium is defined in Appendix B.2.

Government The government provides grants for college education including Pell, runs the federal student loan program with subsidized and unsubsidized loans, funds Social Security, and faces an exogenous government consumption set as a fraction g of gross domestic product (GDP).

Pell grants provide an amount that is based on the expected family contribution, f, and the cost of attendance, $(\kappa + \bar{c})$, where κ denotes tuition and fees and $\bar{c}$ denotes an amount for room and board. The maximum Pell amount is θ^{Pell}_{max}. The Pell grant function is given by

$$\theta^{Pell}(f) = \max[\min[\kappa + \bar{c} - f, \theta^{Pell}_{max} - f], 0] \tag{1}$$

The federal student loan program is characterized by a cumulative student loan limit for subsidized and unsubsidized loans, $\bar{A}$, a cumulative student loan limit for subsidized loans, $\bar{A}_s$, and a student loan interest rate, $r_{SL} = r + \tau_{SL}$, where r is the risk-free interest rate on savings and τ_{SL} is an add-on set by the government. The subsidized loan amount is based on the year of college, j, expected family contribution, and cost of attendance with a cumulative limit given by $\bar{A}_s$. The subsidized loan function is given by

$$a_s(j, a', f) = -\mathbb{I}_{a'<0} \min[-a', \bar{A}_s\left(\tfrac{j}{4}\right), j\max[\kappa + \bar{c} - \theta^{Pell}(f) - f, 0]] \tag{2}$$

The expected family contribution is computed based on parental income. If parental income is less than or equal to the zero EFC income threshold, $y_{EFC=0}$, or income tax is less than 0, then the expected family contribution is 0. Otherwise, it is computed using the following schedule with adjusted available income, $y_{adj}(y) = y - T(y)$, where $T(y)$ is the income tax function:

$$EFC(y) \ = \bar{f}_i + \tau_{f,i}[y_{adj}(y) - \underline{y}_{f,i}] \quad \text{if} \quad y_{f,i} < y_{adj} \leq y_{f,i+1} \tag{3}$$

for $i = 1, ..., n_f$.

Government expenditure is financed with tax revenue collected from a flat consumption tax, τ_c, and a progressive income tax, $T(y)$, levied on pretax income, y. The income tax function follows the specification of Heathcote, Storesletten, and Violante (2017) and is given by

$$T(y) = y - \gamma y^{1-\tau_p} \tag{4}$$

where τ_p governs the tax progressivity and γ is used to balance the government budget constraint in every period as shown in equation (24) in Appendix B.2.

Final goods firm Output is produced by a final goods firm. The production function that combines aggregate capital, K, and aggregate labor, L, is Cobb-Douglas with capital share α. Aggregate labor, in turn, is a CES aggregator of efficiency units of labor with low education, L_ℓ, and high education, L_h, with elasticity of substitution $1/(1-\iota)$ and share parameter ν. The final goods production function is therefore given by:

$$Y = K^\alpha \left(Z \left((\nu L_\ell^\iota + (1-\nu) L_h^\iota)^{1/\iota} \right) \right)^{1-\alpha} \tag{5}$$

where Z is aggregate labor productivity. The capital stock depreciates at rate δ.

Consumers Let j denote the age of consumers; consumers start making decisions when they turn 18 at $j = 1$. Figure 1 illustrates the phases of the consumer's adult life cycle.

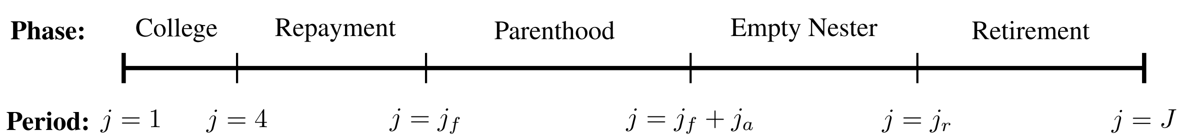


Figure 1: Phases of the consumer's life cycle

Let s denote the skill endowment, which partly indexes the consumer; at the start of adulthood, the consumer is also indexed by a stochastic idiosyncratic earnings productivity, η, initial net assets, a, expected family contribution f, and an indicator of whether the consumer has mistaken beliefs about eligibility for need-based federal financial aid, b.

The skill endowment of a new adult is drawn once from a uniform distribution; it indexes the college enrollment option shock, the exogenous probability of being allowed to continue in college given enrollment, grants, the deterministic life cycle component of earnings, and Social Security transfers. The idiosyncratic stochastic component of earnings, η, follows a lag-1 auto-regressive, or AR(1), process that depends on completed education. Net assets at the start of adulthood are determined by a one-time inter vivos transfer from the consumer's parent. The expected family contribution is a function of parental income. The belief about need-based financial aid eligibility is a one-time draw at age 18, where $b = 1$ indexes a consumer who believes that they are not eligible for any need-based federal financial aid, which may or may not be correct, and $b = 0$ indexes a consumer whose beliefs are based on the true eligibility criterion.[7] Consumers with mistaken beliefs learn about need-based eligibility immediately after enrollment.

With probability $q(s)$, 18-year-old consumers may choose whether to enroll in college; earning a college degree requires four years of completed enrollment. The probability of being allowed to continue from year j to year $j+1$ of college is given by the skill-specific probability $p(s)$.

[7]Note that a consumer with $b = 1$ will have the correct beliefs if they are in fact not eligible for any need-based aid based on the eligibility criterion.

Education is recorded with e; a college student or college graduate has a high level of education, indicated with $e = h$. If the consumer never enrolls, or drops out of college, then e is set to ℓ to reflect their low education level. College enrollees have access to federal student loans, where the stock of debt is recorded with $a < 0$.

A college degree indexes the return to labor supply, w_e, the deterministic component of life cycle earnings, the parameters for the AR(1) process for earnings, and Social Security transfers. While enrolled, college students can only work part-time and incur an effort cost net of consumption value, λ. College expenses include only tuition and fees, κ; these expenses can can be financed with federal student loans, inter vivos transfers from parents, earnings from part-time work while enrolled, and Pell grants, other public grants, and private grants (the latter two are represented by $\theta^{other}(s)$ and $\theta^{pr}(s)$, respectively).

Federal student debt is the only form of debt. Repayment begins at age $j = 5$; subsidized federal loans have interest assessed starting at this age, whereas unsubsidized federal loans are assessed interest for all years of borrowing. During the repayment phase, consumers choose whether to make payments, $\rho_R(j, a)$, on their federal student loans. Those who do not make payments for student loans are considered delinquent and must pay $\rho_D(j, a, y)$ instead. In the period of delinquency, debtors also incur a collection fee and a utility cost. For federal loans, the full payment function $\rho_R(j, a)$ is given by

$$\rho_R(j,a) = \begin{cases} -\left[\dfrac{r_{SL}}{1-(1+r_{SL})^{-(T_{SL}+5-j)}}\mathbb{I}_{j\in(4,T_{SL}+4]} + (1+r_{SL})\mathbb{I}_{j>T_{SL}+4}\right]a & \text{if } a < 0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

If there is an outstanding balance and j is still within T_{SL} periods of the college phase, then the loan is amortized with an interest rate of r_{SL}; otherwise, the outstanding principal plus interest is due. If there is no outstanding loan balance, the payment amount is zero. If a consumer chooses delinquency instead of repayment, their disposable income above the amount $\bar{y}$ is garnished at the rate τ_g, with the partial payment bounded above at the full payment amount $\rho_R(j, a)$:

$$\rho_D(j,a,y) = \min[\tau_g \max[y - T(y) - \bar{y}, 0], \rho_R(j,a)] \tag{7}$$

Upon paying off student loans, consumers solve a consumption-savings problem which at certain ages is affected by the presence of a child. All consumers have a child at age j_f. At the beginning of age $j_f + j_a$, the child becomes an independent agent and each parent altruistically makes an inter vivos transfer to their child after observing the child's skill, s_c, the expected family contribution, f, and the subjective belief about the child's eligibility for need-based financial aid, b.

Consumers retire at age j_r and receive Social Security transfers. Furthermore, consumers survive each period with probability ψ_j, and live for a maximum of J periods. Any assets held at death are redistributed to living consumers as accidental bequests, Tr_j.

Pretax income $y_{j,e,s,\eta,a}$ is determined by age, education, skill, stochastic earnings productivity, and net assets, summarized by the tuple (j, e, s, η, a), as given by

$$y_{j,e,s,\eta,a} = [w_\ell \epsilon_{j,\ell,s} \ell_{pt} \mathbb{I}_{j\leq 4} \mathbb{I}_{e=h} + w_e \epsilon_{j,e,s} \mathbb{I}_{j>4 \text{ or } e=\ell} \mathbb{I}_{j<j_r}]\eta + ss_{e,s} \mathbb{I}_{j\geq j_r} + r\left[a\mathbb{I}_{j>1}\mathbb{I}_{a>0} + Tr_j\right] \quad (8)$$

where w_e is the wage rate given e, $\epsilon_{j,e,s}$ is a deterministic life cycle component, ℓ_{pt} is part-time hours, and $ss_{e,s}$ is the Social Security transfer defined in equation (22) in Appendix B.2.[8]

Consumer utility $U(c, j, e)$ exhibits constant relative risk aversion over household consumption, c, which is split with any cohabiting children according to an adult equivalence parameter ζ_j:

$$U(c, j, e) = \frac{\left(\frac{c}{\zeta_j}\right)^{1-\sigma}}{1-\sigma} - \lambda \mathbb{I}_{e=h \text{ and } j\leq 4} \quad (9)$$

Consumer value functions during the college phase Given their type, (s, η, a, f, b), an 18-year-old (when $j = 1$) decides whether to enroll by solving:

$$\begin{aligned} W(s, \eta, a, f, b) =& (1 - q(s))V(j, \ell, s, \eta, a) \\ &+ q(s)[\max_{d_e\in\{0,1\}} (1 - d_e)V(j, \ell, s, \eta, a) + d_e V(j, h, s, \eta, a, f, b)] \end{aligned} \quad (10)$$

where $V(j, \ell, s, \eta, a)$ is the value of not going to college and $V(j, h, s, \eta, a, f, b)$ is the value of college. The value of not going to college or dropping out is given by

$$\begin{aligned} V(j, \ell, s, \eta, a) &= \max_{c\geq 0, a'} U(c, j, \ell) + \beta\psi_j E_{\eta'|\ell,\eta} V(j + 1, \ell, s, \eta', a') \\ &s.t. \\ &(1 + \tau_c)c + a' = y_{j,\ell,s,\eta,a} + a + Tr_j - T(y_{j,\ell,s,\eta,a}) + \mathbb{I}_{a<0} r_{SL} a \\ &a' \geq \min[a, 0] \end{aligned} \quad (11)$$

where a' is the stock of assets or federal student loans in the next period and β is the discount

[8]The indicator $\mathbb{I}_{j>1}\mathbb{I}_{a>0}$ implies that interest income on the inter vivos transfer accrues to the parents and not the newly emancipated child aged $j = 1$.

factor. The subjective value of enrolling in college for $j = 1, ..., 3$ is given by

$$V(j,h,s,\eta,a,f,b) = \max_{c\geq 0,a'} U(c,j,h) + \beta\psi_j E_{\eta'|\ell,\eta} \tag{12}$$
$$\left[p(s)\max[V(j+1,h,s,\eta',a',f,b),V(j+1,\ell,s,\eta',a')] + (1-p(s))V(j+1,\ell,s,\eta',a')\right]$$
$$s.t.$$
$$(1+\tau_c)c + a' + (1-\theta^{other}(s)-\theta^{pr}(s))\kappa = y_{j,h,s,\eta,a} + a + Tr_j - T(y_{j,h,s,\eta,a})$$
$$+ (1-b)\theta^{Pell}(f) + r_{SL}(\mathbb{I}_{a<0}a - (1-b)a_s(j,a',f))$$
$$a' \geq -\bar{A}\left(\tfrac{j}{4}\right)$$

where $\theta^{Pell}(f)$ and $a_s(j,a',f)$ are functions that determine Pell grants and subsidized federal student loans, defined in equations (1) and (2). The continuation value reflects that consumers may choose to drop out before the start of the next academic year. For consumers with $b = 1$, we assume that they learn the true eligibility criterion immediately after enrolling in college. Therefore, for these consumers, only the enrollment choice is based on their mistaken beliefs. However, conditional on enrollment, the consumption-savings and dropout decisions are based on the case in which $b = 0$. Of course, these consumers do not foresee that they will update their mistaken beliefs when making the college enrollment choice, which is reflected in the continuation value of equation (12). For $j = 4$, the value function is slightly modified from equation (12): the AR(1) draw in the next period will be made from the distribution for the high-education labor and there will be no endogenous dropout decision because the consumer would have graduated in the next period.

5 Parameterization and Model Validation

This section explains the model parameterization in subsection 5.1 and validates the model's calibrated baseline in subsection 5.2.

5.1 Parameterization

Table 3 reports internally calibrated parameters, while Table 10 in the Appendix of the main text reports the externally estimated parameters. In Table 3, the first column contains the parameter symbol; the second column, the parameter description; and the third column, the parameter value. Columns 4 through 6 contain the target moment's description, the moment in the data, and the moment in the calibrated model, respectively. Although parameters and moments are grouped using the most significant one-to-one relationship between each parameter and target moment, and are discussed accordingly, note that the parameters are calibrated jointly and each parameter can affect all target moments.

Panel A of Table 3 presents parameters governed by moments from the HSLS:09. First is $p(s)$, which determines the probability of being allowed to continue in college and is chosen to target persistence rates to the end of the third academic year (Y3), given enrollment in a four-year degree (Y1). Second is λ, the college effort cost net of the consumption value of college, which is set to match observed college enrollment rates in the fall of 2013. Third is $q(s)$, the college option shock, chosen to target enrollment rates for the top family income tercile for each skill tercile; focusing on enrollment rates of young people from high-income families minimizes the role of financial constraints in the enrollment decision. Target moments for these three rows are reported in Table A6 of Appendix A.1.2. Fourth is π_b, the share of 18-year-olds who believe they are not eligible for need-based financial aid regardless of the eligibility criterion (i.e., consumers for whom $b = 1$). This share parameter is set so that the model matches the share of non-enrollees with mistaken beliefs about eligibility reported in Table 2 of Section 3.

Table 3: Internally calibrated parameters

Symbol	Parameter description	Parameter value	Moment description	Data moment	Model moment
Panel A: Moments from the HSLS:09					
$p(s)$	Continuation prob. average	(0.758,0.902,0.955)	Persist to Y3 | Enrolled Y1	*(0.575,0.801,0.912)*	*(0.575,0.803,0.912)*
λ	Net college effort cost	-0.983	Enrolled fall 2013	*0.453*	*0.453*
$q(s)$	Enrollment option shock	(0.614,0.674,0.848)	Enr. fall 2013 | High family (parent) inc.	*(0.251,0.638,0.848)*	*(0.250,0.638,0.848)*
π_b	Share of 18-year-olds with $b = 1$	0.104	Mistaken about eligibility | Non-enrollee	*0.101*	*0.101*
Panel B: Moments from other sources					
β_c	Parent altruism toward child	0.191	Average transfer, normalized	*0.579*	*0.580*
$\bar{c}$	College room and board	0.146	Room + board, normalized	*0.146*	*0.146*
κ	Annual tuition	0.183	Net tuition + fees, normalized	*0.097*	*0.097*
$\bar{y}$	Garnishment-exempt income	0.152	Exempt earnings, normalized	*0.152*	*0.152*
ξ_D	Federal delinquency cost	0.182	Federal delinquency rate	*0.088*	*0.086*
ν	Low-education labor share	0.467	College wage premium | s_2	*1.420*	*1.421*
Z	Aggregate labor productivity	0.573	GDP per capita 18+	*1.000*	*1.000*
β	Discount factor	0.976	Capital-to-output ratio	*3.000*	*3.001*
χ	SS replacement rate	0.188	SS expenditure, fraction of GDP	*0.048*	*0.048*

Panel B of Table 3 reports parameters that are governed by moments from sources other than the HSLS:09. In several rows within this panel, we note that the moment is normalized by GDP per capita for those 18 and over; this value is computed for each year by combining information on GDP from 2013-2015 from BEA (2022, T1.1.5) and population levels from the US Census Bureau found in Census Bureau of the United States (2020). First, the degree of parental altruism, β_c, is set so that the model matches average parent-to-child transfers, as reported in Table A22 of Appendix A.2.3. The cost of college room and board, $\bar{c}$, is set using the average annual value for room and board, and annual tuition, κ, targets average net tuition and fees; both empirical moments are for bachelor's degree programs from 2013-2015 using data contained in a College Board report (Ma, Pender, and Libassi, 2020), supplemented with information from NCES (2019). The income exempt from garnishment in delinquency, $\bar{y}$, is set to 0.152 based on our calculations using results from Yannelis (2020). The parameter governing the costs of being delinquent on public loans, ξ_D, targets the average cohort delinquency rate from 2013 to 2015 reported in FSA (2021b), where the definition of delinquency in the data is a delay in payment of 270 days or more. The parameter that determines the labor share for low-education labor, ν, is set so that the college wage premium for the middle skill tercile matches that observed in the data, reported in Table A21 of Appendix A.2.2. Aggregate labor productivity, Z, is set so that GDP per capita for the population aged 18 and over is 1 in the model. The discount factor, β, is calibrated to target a capital-to-output ratio of 3, consistent with Jones (2016). The Social Security replacement rate, χ, targets the average ratio of total Social Security expenditure to GDP from 2013 to 2015, estimated using data from BEA (2022, T2.1) and BEA (2022, T1.1.5).

5.2 Model validation

We validate the model's calibrated baseline against 5 sets of untargeted moments from the data.

1. Enrollment response to $1,000 subsidy Table 4 presents enrollment response estimates from a quasi-natural experiment in which prospective college students are given a $1,000 subsidy to attend college. The empirical estimate for the enrollment response based on Deming and Dynarski (2009) is 4.00 percentage points. In the model, the analogous non-targeted estimate is 3.16 percentage points, which indicates that the model performs well in comparison to the data.

Table 4: Enrollment response to $1,000 subsidy

Quasi-Natural Experiment	**Data**	**Model**
Enrollment change due to additional $1,000 tuition subsidy	4.00	3.16

Notes: Table 4 presents estimates for the enrollment response given a subsidy of $1,000 in the data and model. Source: Deming and Dynarski (2009).

2. Under-enrollment. Table 5 compares our empirical estimate of under-enrollment with the model's baseline counterpart. In our baseline economy, we compute under-enrollment as the difference between equilibrium enrollment choices and predicted choices made with correct beliefs. That is, a high school graduate is counted as an under-enrollee if they do not enroll with mistaken beliefs, but would enroll otherwise, in partial equilibrium. Under-enrollment among non-enrollees in the model is 1.1 percent. Is this a plausible estimate? We argue yes. In our empirical estimation of mistaken beliefs, we find that 4 percent of non-enrollees do not file for a FAFSA because of only mistaken information about aid eligibility and no other reported reason (Table A12, Panel C in Appendix A.1.4). If we assume that this population will choose to enroll in college with corrected beliefs, then extent of under-enrollment in the data is 4 percent. This is a reasonable interpretative assumption because this population could have cited lack of intent to attend a post-secondary institution as a reason to not file a FAFSA, but they do not.

Table 5: Under-enrollment

Variable	Data	Model
Share under-enrollment | Non-enrollees	4.0	1.1

Notes: Table 5 presents the extent of "under-enrollment" in the data and model. See the main text for the definition of under-enrollment. Source: HSLS:09.

3. College wage premiums by skill Table 6 presents the college wage premium by skill tercile, comparing the model to data. The data estimates are computed using the NLSY97, as reported in Table A21 of Appendix A.2.2. Only the middle tercile's premium is targeted (in italics). In both the data and the model, the college wage premium is computed as the median earnings of individuals with a four-year college degree divided by the median earnings of those without such a degree, specifically for workers aged 25 to 39 within their respective skill levels (the ages are chosen to match the NLSY97 sample). While the wage premium for the middle skill tercile was a calibration target, the model performs well in capturing college wage premiums across all skill terciles. Notably, the college wage premium increases with skill level.

Table 6: College wage premium by skill quantile

Skill	Data	Model
1	1.32	1.32
2	*1.42*	*1.42*
3	1.57	1.55

Notes: Table 6 presents the college wage premium by skill tercile in the data and model. Source: NLSY97.

4. Enrollment rates by parental income and skill Table 7 reports enrollment rates by parental income and young adult skill tercile; data values are reported in Table A3 of Appendix A.1.1. Only enrollment for the top parental income tercile are targeted (in italics). The model generates the same broad qualitative pattern observed in the data: enrollment rates increase in parental income controlling for skill and increase in skill controlling for parental income. However, quantitatively the model understates enrollment for bins outside the top skill and income terciles.

5. Uptake of need-based aid Table 8 reports the share of first year enrollees receiving Pell grants and subsidized student loans, as well as the level of each type of aid conditional on receipt, in the data and model. Data moments are from Table A16 of Appendix A.1.5. While the model does well in the intensive margin, it understates the extensive margin of uptake. In this sense, our estimates

Table 7: College enrollment rates by parental income and youth skill tercile

	Skill Tercile					
	1	2	3	1	2	3
Parental income tercile		**Data**			**Model**	
1	7.2	31.3	61.3	0.6	27.8	79.1
2	15.0	45.4	71.2	2.0	40.6	84.7
3	*25.1*	*63.8*	*84.8*	*25.0*	*63.8*	*84.8*

Notes: Table 7 presents bachelor's degree enrollment rates by family income tercile and youth skill tercile. Source: HSLS:09.

of the welfare cost of under-utilization of need-based aid in the next section are conservative.

Table 8: Need-based financial aid uptake data and model

Aid type	**Margin**	**Unit**	**Data**	**Model**
Pell grant	Extensive	Share of first-year enrollees	33	17
	Intensive	Dollars | Pell grant > 0	$4,413	$3,097
Subsidized student loan	Extensive	Share of first-year enrollees	47	16
	Intensive	Dollars | Pell grant > 0	$3,221	$3,862

Notes: Table 8 presents the extensive and intensive margin of need-based aid uptake of first year enrollees in the data and model. Source: HSLS:09.

6 Main Experiment: Information Intervention

This section discusses the results of our main experiment: an information intervention that eliminates misinformation about need-based aid eligibility. Before moving onto the welfare implications, we analyze the effects for education and skill statistics, macroeconomic aggregates, and prices.[9]

The effects of the information intervention on the model's steady state equilibrium are shown in Table 9. Effects on the model economy are summarized by changes in education and skill statistics (Panel A), macroeconomic aggregates (Panel B), and prices, income tax rate, and transfers (Panel C).

The first row of Panel A reports changes in enrollment rates by skill, which increases especially for the medium- and high-skill. When 18-year-olds who are misinformed learn that they are eligible for means-tested federal aid, that is more likely to change the enrollment decision for those with medium- and high-skill in comparison to those with low-skill because the expected benefits of a

[9] When we measure welfare, the policy maker computes lifetime utilities without the misinformation about need-based grant eligibility, taking as given the policy functions of the consumer.

college degree are smaller for the latter. The next row reports changes in under-enrollment rates. By construction, in the new equilibrium, under-enrollment—which is the difference between equilibrium enrollment choices and predicted choices made with correct beliefs—goes to zero. That statistic is highest for the highest-skilled students in the baseline equilibrium, so under-enrollment changes the most for those with the highest skill. The next row of Panel A indicates that the information intervention barely affects the college graduation rate; this value increases because the skill composition of enrollees shifts towards the highest skill tercile. in the new steady-state, slightly higher enrollment along with a slightly higher graduation rate leads to more college graduates are a share of the population.

Table 9: Steady state changes after information intervention

Category	Variable	Changes from baseline equ.
Panel A: Education and skill statistics	College enrollment rate by s	(-0.35,0.43,0.73)
Units: percentage point change	Underenrollment rate by s	(-0.06,-1.60,-4.47)
	Graduation rate	0.11
	Population share college graduates	0.25
Panel B: Macroeconomic aggregates	Low-education labor (efficiency units)	-0.38
Units: percentage change	High-education labor (efficiency units)	0.77
	Labor	0.12
	Capital	0.09
	Output	0.11
	Consumption	0.12
Panel C: Prices, income tax rate, transfers	Risk-free savings interest rate	0.00
Units: percentage point/percentage change	Wage rate for low-education	0.09
	Wage rate for high-education	-0.13
	Income tax rate | Baseline mean income	-0.02
	Inter vivos transfers	-0.45
	Accidental bequests	0.06
	$\text{ss}_{\ell,s}$ by s	(0.10, 0.10, 0.10)
	$\text{ss}_{h,s}$ by s	(0.03, 0.03, 0.01)

Notes: Table 9 provides results from a steady state comparison of the baseline economy with an economy in which an information intervention eliminates misinformation. Panels A, B, and C report changes in education and skill statistics, macroeconomic aggregates, and prices, income tax rate, and transfers, respectively. Statistics that vary over s are presented as a tuple in the order (s_1, s_2, s_3).

Moving to Panel B, the increase in the mass of college graduates increases the total efficiency units of labor, which increases total labor earnings. Higher earnings, in turn, increase both savings and aggregate capital. This rise in factor inputs increases output and, consequently, increases consumption.

Panel C of Table 9 indicates that the risk-free savings rate barely changes. The wage rate for those with low-education increases, whereas the wage rate for those with high-education decreases. This happens because of the fall and the rise in the labor supply of each respective education group. The next row indicates that the income tax rate barely changes. Inter vivos transfers increase because the economy is richer due to the presence of more college graduates. The changes in

Social Security for the low-education group reflect the change in their wage rate. For the high-education group, there are more college enrollees with higher productivity, which increases Social Security transfers slightly.

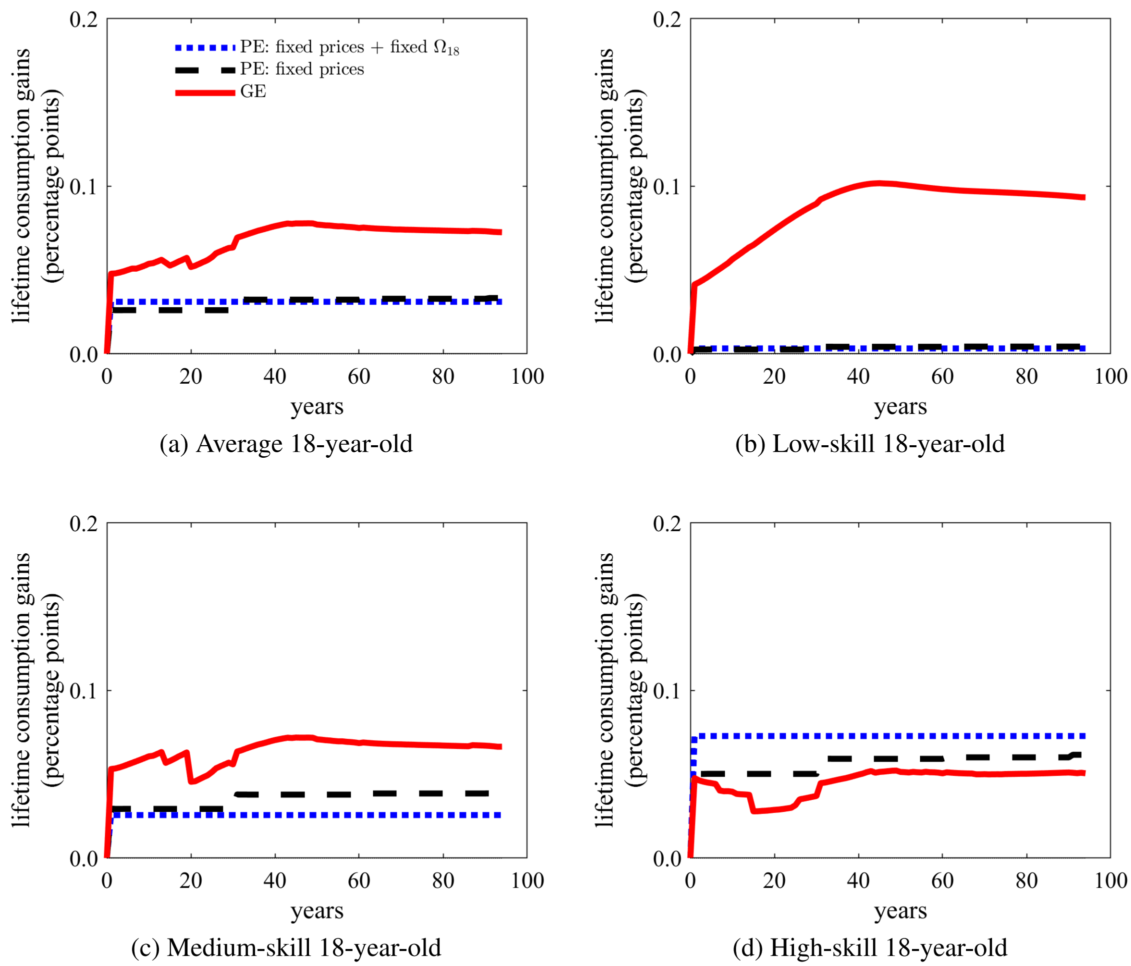
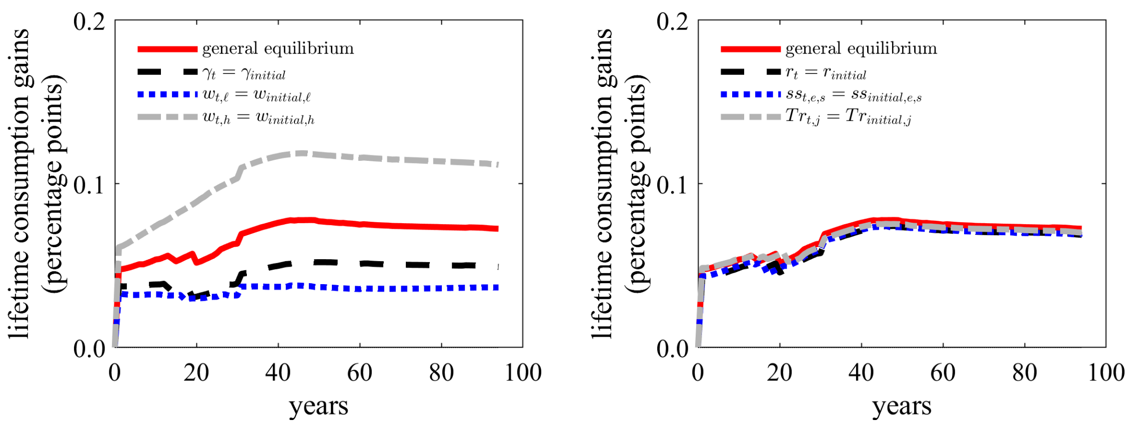


Figure 2: Information intervention: partial and general equilibrium effects

Notes: Figure 2 provides a welfare analysis of the information intervention for 18-year-old consumers in partial and general equilibrium. Subfigures 2a-2d report lifetime consumption gains and losses for the average 18-year-old and the average-18-year-old given skill in each period of the transition path under the following cases separately: (1) a partial equilibrium in which the income tax rate, prices, bequests, Social Security transfers, and the 18-year-old distribution are fixed at their initial steady state values; (2) a partial equilibrium in which the income tax rate, prices, bequests, and Social Security transfers are fixed at their initial steady state values, but the 18-year-old distribution is endogenous; and (3) general equilibrium.

Next, we discuss welfare changes over the transition to the new steady state. In Figure 2 illustrates welfare changes after the information intervention for 18-year-old consumers. In particular, in each subfigure we plot three lines to represent the following cases: (1) a partial equilibrium in which the income tax rate, prices, bequests, Social Security transfers, and the 18-year-old distribution are fixed at their initial steady state values; (2) a partial equilibrium in which the income tax rate, prices, bequests, and Social Security transfers are fixed at their initial steady state values, but the 18-year-old distribution is endogenous; and (3) general equilibrium. Subfigure 2a shows welfare changes for the average 18-year-old; subfigures 2b, 2c, and 2d present the analogous welfare changes by skill endowment (low, medium, and high, respectively). In partial equilibrium, when prices and the distribution of 18-year-olds are fixed, those with the highest skill endowment benefit the most, but the benefits for those with medium- and low-skill are small. Once we take the impact of transfers into account by allowing the distribution of 18-year-olds to change, gains increase for the medium-skill and decrease for the highest skill level. This result highlights the importance of the inter-generational impact on welfare. In general equilibrium, the low- and medium-skill are much better off relative to the initial steady-state, because they benefit from the increase in the wage rate for those with low education. However, the welfare gains of high-skill young people are further dampened in general equilibrium, because the wage rate for those with high education falls due to the rise in the college attainment rate in the population. For the average 18-year-old, both the inter-generational impact and general equilibrium impact act to amplify the gains that stem from the information intervention.



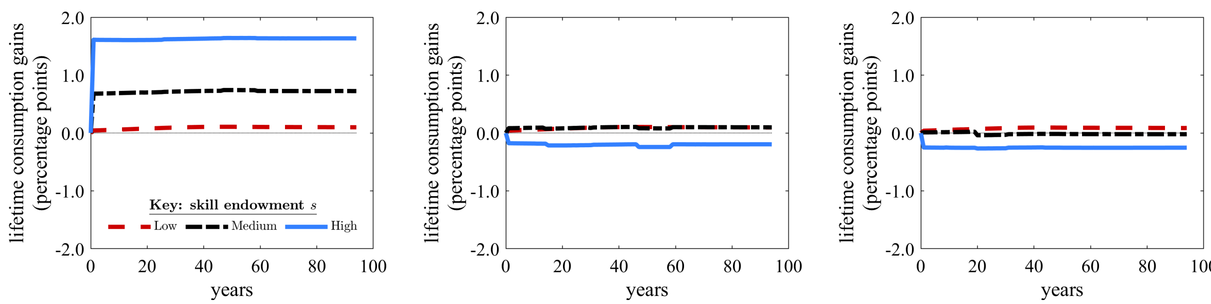
(a) Group 1

(b) Group 2

Figure 3: Decomposing general equilibrium welfare effects of the information intervention

Notes: Figure 3 provides a decomposition of welfare changes due to the information intervention for the average 18-year-old consumer before the college enrollment decision. Subfigures 3a and 3b plot lifetime consumption gains and losses for the average 18-year-old in each period of the transition path under the following cases: general equilibrium, income tax level parameter γ_t fixed at its initial level, wage rate for each education level, $w_{t,e}$, fixed at their initial level, risk-free savings rate, r_t, fixed at its initial level, Social Security transfers, $ss_{t,e,s}$, fixed at their initial level, and accidental bequests, $Tr_{t,j}$, fixed at its initial level. For each partial equilibrium case, while the relevant variable is fixed at its initial level, the other variables change as they do in general equilibrium.

In support of our claim that changes in wage rates drive general equilibrium contributions to welfare changes in Figure 2, Figure 3 decomposes welfare changes for the average 18-year-old by attributing them to each of the several objects that adjust in general equilibrium. For ease of exposition, these objects are divided into two groups based on the magnitude of their contribution to general equilibrium welfare changes. Specifically, subfigure 3a (group 1) plots lifetime consumption changes during the transition for the following partial equilibrium cases: the income tax level parameter, γ_t, and the wage rate for each education group, $w_{t,e}$, each fixed at their initial level while other variables adjust in equilibrium. General equilibrium welfare changes are also included for comparison. Fixing the wage rate at its initial level for those with low education dampens the gains to the average 18-year-old consumer, indicating that the increase in this equilibrium object amplifies the welfare gains. Analogously, fixing the wage rate at its initial level for those with high-education amplifies the welfare gains, indicating that the fall in this equilibrium object dampens the welfare gains. Within this subfigure, the quantitative impact of change in the income tax rate has the smallest effect. Subfigure 3b (group 2) plots welfare changes in partial equilibrium when the risk-free savings rate, Social Security transfers, and accidental bequests are fixed at their initial steady state values, respectively. Their quantitative impact is very small compared to the impact of objects included in group 1.



(a) Low income (b) Middle income (c) High income

Figure 4: Welfare changes by skill and parental income for 18-year-olds with mistaken beliefs

Notes: Figure 4 provides a welfare analysis of the information intervention for 18-year-old consumers who have mistaken beliefs about eligibility in general equilibrium, by parental income and skill.

Figure 4 highlights the heterogeneous effects of the welfare cost of under-utilization. The group that loses the most from misinformation are misinformed 18-year-olds with high-skill from low-income families.

To summarize, the welfare gains from the information intervention that eliminates mistaken beliefs quantifies the welfare costs of these mistaken beliefs in the baseline economy. In partial equilibrium, the costs of misinformation are large for those with high-skill, but the cost to the average 18-year-old is small. However, once we take into account inter-generational and general equilibrium effects more broadly, the costs of misinformation are also larger for the average 18-year-old. Furthermore, under-utilization is most costly for misinformed 18-year-olds with high-skill who come from poor families.

7 Conclusion

Several studies of small-scale information interventions have shown that providing information about college financial aid eligibility has a quantitatively significant effect on college enrollment choices. However, the extent of mistaken beliefs about financial aid eligibility in the aggregate and the welfare costs of this friction have not been quantified. Our study fills this gap in the literature.

Using a public dataset covering high school students in the United States, we estimate the amount of Pell grants that all high school graduates would be eligible for if they were to enroll in a bachelor's degree program, regardless of their actual enrollment outcome. For non-enrollees, we combine our estimates for counterfactual financial aid eligibility with evidence from survey questions about respondent reasons for not filing a FAFSA, and show that 10 percent of non-enrollees do not file a FAFSA and are mistaken about their eligibility for Pell grants. We then quantify the welfare costs of these mistaken beliefs by embedding them into a structural model of college enrollment choice. After calibrating the model, we shut off mistaken beliefs and examine the welfare impact over the transition to the new steady state.

We find that, in partial equilibrium, welfare losses from mistaken beliefs are primarily borne by those with higher skill. In general equilibrium losses are borne across all skill levels, however, and are larger on average. Our findings indicate that large-scale information interventions can be even more beneficial once general equilibrium effects are taken into account. Overall, mistaken beliefs are most costly in general equilibrium for those with high skill, poor parents, and mistaken beliefs; this indicates that the target group from previous experimental work on information interventions benefits the most in a calibrated structural model which implements a similar intervention at a larger scale.

Our approach currently abstracts from fixed costs of applying for aid, and quantifies the mass of mistaken young adults using a conservative definition of being eligible for aid (i.e., using only imputed Pell grant eligibility, rather than any need-based aid). We aim to incorporate fixed costs in future work.

Appendix

Table 10 presents the externally estimated parameters. Panel A governs demographics, which we set by assumption; Panels B uses estimates from the literature to discipline preferences and technologies; Panel C draws on our estimations using the HSLS:09, NLSY97, and PSID, which are explained in more detail in the Online Appendix; and, finally, Panel D uses statutory values for government policy and (in one case) an estimate from the literature.

In Panel A, the fertility period, j_f, is set to 13 so that consumers have a child when they turn 30; the age adulthood begins, j_a, is set to 18; j_r is chosen so that the retirement age is 65; and, finally, J sets maximum life span to 100 years. For $j < j_f + j_a$, we set survival probabilities ψ_j to one to rule out children without parents; ages $j \geq j_f + j_a$ use estimates from the 2010 Social Security Administration Life Tables presented in Bell and Miller (2020).

Table 10: Other externally estimated parameters

Parameter	Description	Data Target	Value
Panel A: Demographics			
j_f	Child bearing age	30 years	13
j_a	Years for child to move out	18 years	18
j_r	Retirement age	65 years	48
J	Maximum life span	100 years	83
ψ_j	Survival probability	2010 SSA Life Tables	-
Panel B: Preferences and technology			
σ	Risk aversion	Chetty (2006)	2
α	Capital share	Kydland and Prescott (1982)	0.360
δ	Depreciation rate	Krueger and Ludwig (2016)	0.076
ι	Elasticity of substitution	Card and Lemieux (2001)	0.800
ζ_j	Adult equivalence scale	OECD modified scale	$1 + 0.3\mathbb{I}_{j_f \leq j < j_f + j_a}$
Panel C: Life cycle earnings profile and hours worked in college			
$\epsilon_{j,e,s}$	Deterministic component	PSID and NLSY97	Table A20, App. A.2.1
ρ_e	AR(1) persistence for $e = (\ell, h)$		$(0.851, 0.886)$
σ_e^2	AR(1) variance for $e = (\ell, h)$		$(0.083, 0.072)$
ℓ_{pt}	Working hours while in college	HSLS:09	0.348
Panel D: Government student aid policy and grants			
$y_{EFC=0}$	Income threshold 0 EFC, normalized	Statutory	0.345
$y_{f,i}$	AAI thresholds for EFC, normalized		(-∞,-0.049,0.220,0.276,0.332,0.389,0.445,∞)
$\bar{f}_i$	Lower-bound for EFC, normalized		(0,0,0.048,0.062,0.079,0.098,0.120)
$\tau_{f,i}$	Marginal rate for EFC, normalized		(0,0.22,0.25,0.29,0.34,0.40,0.47)
$\underline{y}_{f,i}$	Lower income bound for EFC marginal rate, normalized		(0,0,0.220,0.276,0.332,0.389,0.445)
$\bar{A}$	Subsidized and unsubsidized loan limit, normalized		0.377
$\bar{A}_s$	Subsidized loan limit, normalized		0.265
θ_{max}^{Pell}	Pell maximum, normalized		0.079
τ_{SL}	Interest rate add-on		0.021
T_{SL}	Maximum years to repay		10
τ_g	Federal SL garnishment rate		0.150
$\theta^{other}(s)$	Public grants net of Pell by s	HSLS:09, Krueger and Ludwig (2016)	(0.195, 0.230, 0.308)
$\theta^{pr}(s)$	Private grants by s		(0.134, 0.141, 0.157)
ϕ_D	Student loan collection fee	Luo and Mongey (2019)	0.185
Panel E: Government spending and tax policy			
g	Government consumption	BEA	0.147
τ_p	Income tax progressivity	CBO	0.177
τ_c	Consumption tax rate	OECD	0.044

Panel B reports parameters that govern preferences and the goods production technology. It begins with the relative risk aversion parameter, σ, which is set to 2 based on Chetty (2006). The adult equivalence scale, ζ, is set to 0.3 following the Organization for Economic Co-operation and Development (OECD) modified scale. The capital share parameter, α, is set to 0.36 following Kydland and Prescott (1982). The depreciation rate of capital, δ, is set to 0.076, as in Krueger and Ludwig (2016). The parameter that dictates the elasticity of substitution between low- and high-education labor, ι, is set to 0.8, which implies an elasticity of substitution of 5. This value is in the middle of the range (between 4 and 6) reported in Card and Lemieux (2001) after controlling for imperfect substitutability across age groups.

The first three rows of Panel C contain objects that determine life cycle earnings. First is the deterministic component of the life cycle earnings process, $\epsilon_{j,e,s}$; second, the persistence parameter of the AR(1) productivity shock, ρ_e; and, third, the variance of the AR(1) productivity shock, σ_e^2. These objects are estimated using data from both the PSID and NLSY97, with full results presented in Table A20 and details provided in Appendix A.2.1. The last row of Panel C reports working hours while in college, ℓ_{pt}, which is set to the average weekly hours worked as a fraction of 40 (full time) for third-year college students in the HSLS:09, as reported in Table A8 of Appendix A.1.2.

Panel D reports government policy parameters related to student aid, as well as subsidy rates for grants other than Pell grants. The first eight rows of this panel are normalized by GDP per capita for those 18 and over; this value is computed for each year by combining information on GDP from 2013-2015 from BEA (2022, T1.1.5) and population levels from the US Census Bureau found in Census Bureau of the United States (2020). In rows 1-5 we report parameters that govern the EFC function in the model, which are drawn from the EFC formula guide for 2013-2014 prepared by the Federal Student Aid Office, U.S. Department of Education (2014). The first row reports the income threshold below which households are assigned an automatic zero EFC, $y_{EFC=0}$, which is set to 0.345. If the household does not qualify for an automatic zero EFC, then the EFC is computed using a schedule provided in equation (3). The various parameters of this schedule are reported in rows 2-5. Rows 6, 7, and 8 report limits for federal aid: $\bar{A}$ is the net borrowing limit for any federal student loans which is set to 0.377, $\bar{A}_s$ is the net borrowing limit for subsidized loans which is set to 0.265, both determined using the limits for four years of college from Smole (2019), and θ_{max}^{Pell} is the maximum Pell amount an individual can be awarded, set to 0.079 using the amount from the Office of Postsecondary Education, Federal Student Aid (2013). The ninth row contains τ_{SL}, the interest rate add-on set to 0.021 as reported by the Chief Operating Officer for Federal Student Aid (FSA) in Chief Operating Officer for FSA (2021). Row 10 contains the number of years for repayment on a student loan, T_{SL}, which is set to 10 based on the standard repayment plan explained in Smole (2019), and eleventh is the garnishment rate conditional on delinquency on student loans, τ_g, which is set as established by the 2005 Deficit Reduction Act (109th Congress of the United States of America, 2006). The following two rows of Panel D contain the vectors representing shares of tuition and fees paid with grants and scholarships from public sources other than Pell grants, $\theta^{other}(s)$, and private sources, $\theta^{pr}(s)$, which are determined using data from the HSLS:09 incorporating estimates from Krueger and Ludwig (2016); see Appendix A.1.1 for details and Table A7 for results. The last row contains the student loan collection fee, ϕ_D, which is set to 0.185 following Luo and Mongey (2019).

Panel E reports parameters related to government spending and tax policy. Government consumption as a share of GDP, g, is set to 0.141 using estimates of the numerator and denominator from the Bureau of Economic Analysis (BEA) in BEA (2022, T1.1.5) and BEA (2022, T3.1). The income tax progressivity, τ_p, is set to 0.177 following our estimation results using data from the Congressional Budget Office (CBO) provided in U.S. Congressional Budget Office (2018a,b), with the estimation procedure described in Appendix A.3 and point estimates presented in Table A24. Finally, we estimate the consumption tax rate τ_c to be 0.044 by applying the method of Mendoza, Razin, and Tesar (1994) to OECD data for the period 2013-2015 (OECD, 2024c,b,a); estimation results are presented in Table A26 of Appendix A.4.

References

109th Congress of the United States of America (2006). Deficit Reducation Act of 2005 (Public Law 109-171). (accessed 05.10.2024).

Bell, F. C. and M. L. Miller (2020). Life Tables for the United States Social Security Area, 1900-2100, Actuarial Study No. 6, Table 6. (accessed 06.01.2022).

Board of Governors of the Federal Reserve System (2014). Changes in U.S. Family Finances from 2010 to 2013: Evidence from the Survey of Consumer Finances. *Federal Reserve Bulletin 100*(4).

Bureau of Economic Analysis, U.S. Department of Commerce (2022a). Table 2.1. (accessed 05.15.2022).

Bureau of Economic Analysis, U.S. Department of Commerce (2022b). Table 1.1.5. (accessed 05.15.2022).

Bureau of Economic Analysis, U.S. Department of Commerce (2022c). Table 3.1. (accessed 05.15.2022).

Card, D. and T. Lemieux (2001). Can falling supply explain the rising return to colleg for younger men? A cohort-based analysis. *The Quarterly Journal of Economics 116*(2), 705–746.

Census Bureau of the United States (2000). National Intercensal Tables: 1990-2000. Produced by the Census Bureau of the United States. (accessed 05.23.2024).

Census Bureau of the United States (2010). National Intercensal Tables: 2000-2010. Produced by the Census Bureau of the United States. (accessed 05.23.2024).

Census Bureau of the United States (2013). Table HINC-06. Income Distribution to $250,000 or More for Households: 2012. Produced by the Census Bureau of the United States using the Current Population Survey, Annual Socioeconomic Supplement. (accessed 02.11.2024).

Census Bureau of the United States (2020). National Population by Characteristics: 2010-2019. Produced by the Census Bureau of the United States. (accessed 05.23.2024).

Chetty, R. (2006). A New Method of Estimating Risk Aversion. *The American Economic Review 96*(5), 1821–1834.

Chief Operating Officer for Federal Student Aid (2021). Annual Notice of Interest Rates for Fixed-Rate Federal Student Loans Made Under the William D. Ford Federal Direct Loan Program. *Federal Register 86*(152), 44003–44005. (accessed 07.05.2022).

Collins, B. (2021). Federal Student Aid: Need Analysis Formulas and Expected Family Contribution. (Congressional Research Service of the United States R44503).

Congressional Research Service of the United States [author name redacted] (2018). Campus-Based Student Financial Aid Pograms Under the Higher Education Act. (CRS RL31618).

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

Dortch, C. (2023). Federal Pell Grant Program of the Higher Education Act: Primer. (Congressional Research Service of the United States R45418).

Dynarski, S., C. Libassi, K. Michelmore, and S. Owen (2021). Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Students. *American Economic Review 111*(6), 1721–1756.

Dynarski, S. and J. E. Scott-Clayton (2006). The Cost of Complexity in Federal Student Aid: Lessons from Optimal Tax Theory and Behavioral Economics. *National Tax Journal 59*(2).

Federal Student Aid Office, U.S. Department of Education (2014). 2013-2014 Expected Family Contribution (EFC) Formula Guide. (accessed 01.17.2024).

Flood, S., M. King, R. Rodgers, S. Ruggles, J. R. Warren, D. Backman, A. Chen, G. Cooper, S. Richards, M. Shouweiler, and W. Michael (2023). *IPUMS CPS: Version 11.0 Annual Socioeconomic Supplement (ASEC)*. IPUMS (Minneapolis, MN).

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

Hoxby, C. M. and S. Turner (2015). What High-Achieving Low-Income Students Know About College. *American Economic Review: Papers and Proceedings 105*(5), 514–517.

Jones, C. (2016). Facts About Economic Growth. *Handbook of Macroeconomics 2A*.

Krueger, D. and A. Ludwig (2016). On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium. *Journal of Monetary Economics 77*, 72–98.

Kydland, F. E. and E. C. Prescott (1982). Time to Build and Aggregate Fluctuations. *Econometrica 50(6)*, 1345–1370.

Lin, D. (2022). Methods and Assumptions of the CPS ASEC Tax Model. *SEHSD Working Paper FY-2022-18*.

Luo, M. and S. Mongey (2019). Assets and job choice: Student debt, wages and amenities. *Unpublished manuscript*.

Ma, J., M. Pender, and C. Libassi (2020). Trends in College Pricing and Student Aid 2020. *College Board*.

Mendoza, E. G., A. Razin, and L. L. Tesar (1994). Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption. *Journal of Monetary Economics 34*, 297–323.

Moschini, E. G., G. Raveendranathan, and M. Xu (2023). Optimism About Graduation and College Financial Aid. *Unpublished manuscript*.

NASFAA Monograph, 8th Edition (2023). Developing the cost of attendance: Practical information for student aid professionals. (accessed 1.26.2024).

National Center for Education Statistics, U.S. Department of Education (2017). Table 330.10. Produced and distributed by the Institute of Education Sciences, National Center for Education Statistics (NCES). (accessed 01.04.2023).

National Center for Education Statistics, U.S. Department of Education (2019). Table 330.10. Produced and distributed by the Institute of Education Sciences, National Center for Education Statistics (NCES). (accessed 01.04.2022).

National Center for Education Statistics, U.S. Department of Education (2020). High School Longitudinal Study of 2009 (public use data file). Produced and distributed by the Institute of Education Sciences, National Center for Education Statistics (NCES). (accessed 10.11.2022).

National Center for Education Statistics, U.S. Department of Education (2022). Table 302.10. Produced and distributed by the Institute of Education Sciences, National Center for Education Statistics (NCES) (accessed 01.03.2023).

OECD (2024a). Annual government expenditure by function (COFOG). Public use dataset, OECD.

OECD (2024b). Annual household final consumption expenditure on the territory and abroad. Public use dataset, OECD.

OECD (2024c). Details of Tax Revenue - United States. Public use dataset, OECD.

Office of Postsecondary Education, Federal Student Aid (2013). 2013-2014 Federal Pell Grant Payment and Disbursement Schedules. *Dear Colleague Letter GEN-13-06*. (accessed 1.26.2024).

Office of Postsecondary Education, Federal Student Aid (2022). FAFSA Simplification Act Changes for Implementation in 2023-24. *Dear Colleague Letter GEN-22-15*. (accessed 1.26.2024).

Program Communications Division, Federal Student Aid (2013). Federal Student Aid Handbook 2013-2014. Technical handbook, U.S. Department of Education.

Smole, D. P. (2019). Federal Student Loans Made Through the William D. Ford Federal Direct Loan Program: Terms and Conditions for Borrowers. (Congressional Research Service of the United States R45931).

U.S. Bureau of Labor Statistics and U.S. Census Bureau (2006). Design and Methodology: Current Population Survey. Technical report 66, U.S. Bureau of Labor Statistics and U.S. Census Bureau.

U.S. Congressional Budget Office (2018a). The Distribution of Household Income, 2014.

U.S. Congressional Budget Office (2018b). The Distribution of Household Income, 2015.

U.S. Department of Education, Office of Federal Student Aid (2021). National Default Rate Briefing for FY 2018 Official Cohort Default Rates. (accessed 05.15.2022).

Yannelis, C. (2020). Strategic default on student loans. *Unpublished manuscript*.

Online Appendix for:
"Mistaken Beliefs About Eligibility for College Financial Aid:
Empirical Evidence and General Equilibrium Impact"

by Emily G. Moschini and Gajendran Raveendranathan

Contents

A Data Appendix

A.1 The HSLS:09

We use the The High School Longitudinal Study of 2009 (HSLS:09) as our main data source in this paper.

A.1.1 Postsecondary outcomes

Assigning postsecondary outcomes to high school graduates As part of our sample cleaning procedure we assign postsecondary outcomes to the set of high school graduates and drop observations for whom this assignment is not possible or which have inconsistent records. To assign post-secondary outcomes for HSLS:09 cleaned sample members in the fall of 2013 and the second follow-up collected in 2016, we proceed as follows. First, we recognize that student records collected from postsecondary institutions are likely to be more reliable measures of postsecondary enrollment than survey responses. Because of this, we prioritize information on postsecondary outcomes collected from student records submitted by postsecondary institutions to the HSLS:09. Specifically, we begin by identifying high school graduates who enroll in postsecondary education for the first time in the fall of 2013 (the 2014 academic year). We flag this group by identifying those who have a nonmissing value for at least one of the variables X4PS1START, X5PFYEAR, or X5POSTHSAY that indicates enrollment in the academic year of interest if it is not missing. We also require that these variables do not contradict each other if more than one of them is non-missing. For this group of enrollees, we use the variable X5PFYDEGREE (a variable from the student records/postsecondary transcript data collection) to assign the individual to less than 2-year program, an associate's degree/ 2-year program, or a bachelor's degree/4-year program. For observations that remain unassigned to a postsecondary outcome, we then use X4PS1DEGTYPE1 (a variable from the second follow-up wave of data collection) to allocate observations to the relevant element of the set of possible postsecondary outcomes.

The next step is to clean out those who are identified as "did not enroll" according to this procedure but who have inconsistent observations in some way (a total of 25 observations). We drop observations who are flagged as having enrolled in the fall of 2013 but have no record of what sort of degree they enrolled in. We also drop those who earned their high school credential before or after the 2013 calendar year, or whose high school credenial was something other than a high school diploma (for example, a GED). We discard those who are identified as not enrolling in any degree type but who have a record of attempting between zero and 10 academic credits. We also discard those whose student aid records indicate that they received no aid overall, but who have positive aid values for specific components of aid. Finally, we discard observations had oustanding

Stafford loan or Pell grant amounts by 2013, and whose income is above the 99th percentile of the sample, but whose record indicates they received a subsidized Stafford loan.

Table A1 reports the enrollment rates that result from this procedure and enrollment rates from the National Center for Education Statistics (NCES) for the same period for comparison (National Center for Education Statistics, U.S. Department of Education, 2022). The NCES tabulations draw on the Current Population Survey for the 2013 year of interest. Our procedure yields enrollment rates in the cleaned sample that align well with the national NCES tabulation drawn from a different data source.

Table A1: Postsecondary education enrollment rates in the fall of 2013

	HSLS:09 cleaned sample	**NCES Table 302.10**
Enrollment in fall 2013	**HS graduates spring 2013**	**Recent HS completers**
Did not enroll	0.253	
Less than 2-year	0.079	
Associate's (2-year)	0.214	0.238
Bachelor's (4-year)	0.453	0.421
Observations	7,045	

Notes: Table A1 reports enrollment rates for the cleaned and conditional sample from the HSLS:09 and postsecondary outcomes for recent high school graduates aged 16-24 from the NCES in the last column on the right, for comparison. HSLS:09 weights: second follow-up student longitudinal weights. Source: High School Longitudinal Study of 2009 and NCES Table 302.10 for 2013.

Descriptive statistics of postsecondary outcomes in fall 2013 Table A2 reports statistics conditional on post-secondary outcomes in the 2013-2014 academic year for the cleaned sample of Table 1 in the main text. Panel A indicates that there is sorting by skill into post-secondary outcomes, with higher-GPA young adults enrolling in BA programs, then AA programs, and so on. The annual cost of a BA program is much higher than 2-year and less-than 2-year programs. Panel B of Table A2 reports education outcomes by 2016. These outcomes represent the highest education outcome for the cleaned sample; observations are assigned the outcome they are currently enrolled in or from which they have received a degree. Among those who do not enroll in any postsecondary education in the fall of 2013, about 87 percent remain unenrolled. By contrast, many of those who enroll in a less-than 2-year program or an associate's degree program continue to a higher level of postsecondary education by 2016. Among those who initially enroll in a BA program, about 84 percent persist to their third academic year.

Post-secondary enrollment breakdowns by skill and family income How do high school graduates sort into bachelor's degree programs by skill and family income? Table A3 breaks down enrollment in a bachelor's degree program in the fall of 2013 by family income and high school

Table A2: Descriptive statistics of postsecondary outcomes by fall 2013 enrollment status

Category	**Variable**	**Enrollment status in fall of 2013** **Did not enroll**	**l.t. 2-year**	**AA (2-year)**	**BA (4-year)**
Panel A: as of fall of 2013	High school GPA	2.401	2.811	2.846	3.529
	Annual tuition and fees	.	6,286	3,563	17,687
Panel B: highest of current enr. or degrees received as of 2016	Not enrolled and no degrees	0.873	0.262	0.338	0.091
	Less than 2-year	0.039	0.238	0.034	0.007
	Associate's (AA, 2-year)	0.056	0.281	0.396	0.060
	Bachelor's (BA, 4-year)	0.032	0.218	0.233	0.842
	Observations	1,395	476	1,256	3,918

Notes: Table A2 reports attributes of postsecondary outcomes in the fall of 2013 for the HSLS:09 cleaned sample of Table 1. Cell contents are means. Dollars are in 2012 USD. Weights: second follow-up student longitudinal weights. Source: High School Longitudinal Study of 2009.

GPA. This table shows that enrollment is increasing in both skill and family income; for the highest skill tercile, the gap between the bottom and top family income tercile is more than 20 percent. Table A4 shows that this qualitative pattern remains once enrollment in 2-year associate's degree programs is included in the enrollment frequencies.

As shown in Table A4, enrollment in postsecondary education is increasing in skill and parental (family) income when enrollment in a 2-year program is included, although the gap between the lowest and the highest income terciles is narrowed for a given skill tercile once 2-year programs are included, relative to Table A3.

Table A3: Bachelor's degree program enrollment patterns by skill and parental income

Parental income tercile	**High school GPA tercile** 1	2	3	**All**
1	0.072	0.313	0.613	0.244
	(0.009)	(0.019)	(0.023)	(0.010)
2	0.150	0.454	0.712	0.436
	(0.015)	(0.017)	(0.015)	(0.010)
3	0.251	0.638	0.848	0.680
	(0.022)	(0.016)	(0.009)	(0.009)
All	0.128	0.466	0.765	0.453
	(0.008)	(0.011)	(0.008)	(0.006)
Observations	1,801	2,235	3,009	7,045

Notes: Table A3 reports the share of the cleaned sample from Table 1 who enroll in a bachelor's degree program by family income tercile (rows) and high school GPA tercile (columns). Standard errors are in parentheses. HSLS:09 weights: second follow-up student longitudinal weights. Source: High School Longitudinal Study of 2009 and NCES Table 302.10 for 2013.

As shown in Table A4, enrollment in postsecondary education is increasing in skill and family income when enrollment in a 2-year program is included, although the gap between the lowest and the highest income terciles is narrowed for a given skill tercile once 2-year programs are included, relative to Table A3.

Table A4: Enrollment in either 2- or 4-year programs in fall of 2013

	High school GPA tercile			
Family income tercile	1	2	3	**All**
1	0.332	0.629	0.828	0.514
	(0.017)	(0.020)	(0.018)	(0.012)
2	0.378	0.733	0.878	0.663
	(0.020)	(0.016)	(0.011)	(0.010)
3	0.519	0.811	0.937	0.826
	(0.025)	(0.013)	(0.006)	(0.007)
All	0.379	0.724	0.900	0.668
	(0.011)	(0.009)	(0.005)	(0.006)
Observations	1,801	2,235	3,009	7,045

Notes: Table A4 reports enrollment rates in either a 2-year or a 4-year program in the fall of 2013. Sample: cleaned HSLS:09 sample. Weights: Postsecondary transcripts/student records longitudinal weights. Source: HSLS:09.

Summary statistics with conditional sample for comparison Our model framework represents students who either enroll in a BA or do not enroll in any postsecondary education. Table A5 compares the effect of dropping high school graduates who enroll in postsecondary education other than a bachelor's degree program on sample composition, relative to a cleaned sample which includes these observations. High school GPA increases slightly, as does family income (the share below 185 percent of the poverty correspondibly falls, relative to the cleaned sample). The share with 2 parents rises slightly; household size and number of siblings in postsecondary education are mostly unchanged.

Table A5: Summary statistics

		Sample		
Category	**Variable**	**Raw**	**Cleaned**	**Conditional**
Panel A: Student attributes	High school GPA	2.816	3.040	3.125
	Female	0.506	0.502	0.493
Panel B: Family attributes	Family income	78,496	88,305	94,982
	Below 185 % of poverty line	0.394	0.329	0.307
	Two-parent family	0.873	0.875	0.881
	Household size	4.269	4.248	4.250
	Number siblings in PSE	0.360	0.422	0.425
Panel C: Enrollment outcome in fall 2013	Did not enroll	-	0.253	0.358
	Less than 2-year	-	0.079	0.000
	Associate's (AA, 2-year)	-	0.214	0.000
	Bachelor's (BA, 4-year)	-	0.453	0.642
	Observations	13,283	7,045	5,313

Notes: Table A5 reports means of variables for the raw, cleaned, and conditional sample from the HSLS:09. Weights: second follow-up student longitudinal weights. Source: HSLS:09.

A.1.2 Model parameterization tabulations

Moments used to parameterize the model of the main text and provided below. Table A6 presents moments related to enrollment and persistence rates among high school graduates in the HSLS. These rates are broken down by skill bin, where skill is measured with honors-weighted high school GPA. Table A7 reports, by skill bin, total merit-based and means-tested grants received during the first year of enrollment as a share of tuition and fees overall, then allocated to private or public sources using shares from Krueger and Ludwig (2016), then total Pell grants as a share of tuition and fees. The last column in this table is the subsidy rate of public grants net of the subsidy rate of pell grants. Lastly, Table A8 reports the average hours worked by third-year BA enrollees, as a fraction of full time (40 hours a week).

Table A6: BA Enrollment and persistence by skill bin

	Enr. in BA	Enr. in BA | Highest income tercile	Enr. Y3 | Enr. Y1
1	0.128	0.251	0.575
2	0.466	0.638	0.801
3	0.765	0.848	0.912
Total	0.453	0.680	0.842

Notes: BA enrollment rates and the condition for being enrolled in a BA in Y1 in the third column refer to enrollment in the fall of 2013. Weights: Second Follow-up student longitudinal weights for the first two columns; PETS-SR longitudinal weights for the third column. Source: HSLS:09.

Table A7: Grants subsidy rates by skill bin

	Grants/TF	Prv Grants/TF	Pub Grants/TF	Pell grants/TF	Pub Grants net Pell/TF
1	0.446	0.134	0.312	0.117	0.195
2	0.470	0.141	0.329	0.099	0.230
3	0.524	0.157	0.367	0.059	0.308

Notes: Weights: PETS-SR longitudinal weights. Source: HSLS:09.

Table A8: Hours worked Y3 as a fraction of 40 hours

Ave. hours worked/40	0.348

Notes: Weights: Second follow-up student longitudinal weights. Source: HSLS:09.

A.1.3 Imputation of means-tested financial aid eligibility

We use the HSLS:09 to impute the expected family contribution (EFC), which allows us to assign financial aid from the Pell grant to all high school graduates in our cleaned sample. With this information, we compare imputed values with realized values for students who enroll in a 4-year bachelor's degree program to validate our imputation. We then document patterns of counterfactual aid eligibility among FAFSA non-filers who never enroll in post-secondary education and who exhibit mistaken beliefs about their own eligibility for federal financial aid.

Our main source of discipline on how the attributes of FAFSA filers map into the EFC is the *Expected Family Contribution (EFC) Formula Guide* published by the US Department of Education's Federal Student Aid office for the 2013-2014 academic year (Federal Student Aid Office, U.S. Department of Education, 2014). In particular, we use Worksheet A for dependent students and the associated Tables A1-A7; whenever we cite Worksheet Table AX, where X is a number, we are referring to Tables recorded in Worksheet A of Federal Student Aid Office, U.S. Department of Education (2014). We construct HSLS:09 survey analogs for the inputs into Worksheet A and then operate on them using the parameters provided in Worksheet Tables A1-A7 of the *Expected Family Contribution (EFC) Formula Guide*. We abstract from FAFSAs expected student contribution for dependent students, so that the expected family contribution is equal to the parents' contribution. Table A9 summarizes the mapping from FAFSA formula inputs to observable counterparts that we use to impute financial aid eligibility (as opposed to realized aid receipt).

Beginning with Category I of Table A9, we measure family income in the HSLS:09 using non-imputed values for the categorical variables X1FAMINCOME and X2FAMINCOME, which record discretized family income reported in the base year and first follow-up of the survey, respectively. We drop imputed income observations becaues they worsen the imputation fit to the data. For each variable, we construct income by first assigning dollar values as the midpoint of each category's range; for the top category we assign the value equal to 1.8 times the lower bound for this interval. The factor of 1.8 is arrived at by taking the ratio of the average income for those within the top income bin as reported by the Current Population Survey's Annual Socioeconomic Supplement (CPS ASEC) tabulation of Census Bureau of the United States (2013), which is \$431,783 in 2012 dollars, and the lower bound of the top bracket in the HSLS:09, which is \$235,000 in 2012 dollars. Finally, we use the Consumer Price Index to convert current dollars to 2012 dollars. We then keep the most recent income observed (as opposed to taking the average across observed incomes, for example) because we want an approximation of the input into the FAFSA form, which is the adjusted gross income from the previous year. Taking the average would be more appropriate if we were trying to approximate permanent income or to average out shocks within a given family.

Moving to Category II of Table A9, we construct federal tax liability using our estimation of the parameters summarizing the federal tax system from the CPS ASEC (Flood et al., 2023), imposing the functional form for tax revenue of $y - \lambda y^{1-\tau}$ and estimating both λ and τ. This estimation and its results are described in more detail in Appendix A.1.7. We then evaluate the expression for tax liability setting y equal to income that are observed for a given family in the HSLS:09. We assign state and other tax allowance using the Worksheet Table A1 values for unobserved state of residence—because state of residence is unreported in the public HSLS:09—evaluated using observed family income constructed as explained above. The Social Security tax allowance for each resident working parent is constructed using Worksheet Table A2 and observed income; the income protection allowance uses Worksheet Table A3 and the household size at the point of the first follow-up reported in the HSLS:09 with variable X2HHNUMBER. The employment expense allowance is constructed with information on the number of resident working parents from the HSLS:09 and parameters reported in the FAFSA formula worksheet. Total allowances against income are then defined as the sum of all allowances in Category II.

Category III of Table A9, available income, is the difference between total income and total allowances against income.

Category IV of Table A9, parents' contribution from assets, sets the net worth of all FAFSA filers to the median family net worth for families where the head is between 45 and 54, which is $105,000, taken from Table 2 of the Federal Reserve Board of Governor's Bulletin on the 2013 SCF (Board of Governors of the Federal Reserve System, 2014). The asset protection allowance is taken from Worksheet Table A5 evaluated using HSLS:09 information on the number of resident parents using the variable P2MARSTAT. The parents' contribution from assets is the difference between net worth and the asset protection allowance, multiplied by the asset conversion rate of 0.12 taken from the *Expected Family Contribution (EFC) Formula Guide*.

Category V of Table A9, the adjusted available income (AAI) is the sum of the parents' available income and their contribution from assets. Their contribution from AAI is defined using Worksheet Table A6, and is then divided by the number of children in college in the 2014 academic year, which is assigned in the HSLS:09 using variables P2INCLG2013,P2SIBSTARTCLG, and P2SIBCLGGRAD. The result is the total contribution from parents. When contributions from the student are set to 0, as we assume in our approximation, the parent contribution is also equal to the expected family contribution (EFC) so that our imputation result for the EFC in Category VI of Table A9 is equal to the parent contribution in line # 28.

In order to impute the expected cost of attendance (COA), we compute the observed cost of attendance by starting with observed tuition and fees in the HSLS:09 for those who enroll in a 4-year program and add the cost of room and board for a 4-year program, drawn the NCES National Center for Education Statistics, U.S. Department of Education (2017), when the individual does not live at home. We then assign the imputed COA as the mean realized of cost of attendance for those enrolled in a BA in the fall of 2013, taking the mean by high school GPA and family income quintile assigned using the distribution of high school graduates. This imputation takes into account that students with differing skill and family income levels may be more likely to live at home and builds that into the estimated imputed COA.

Table A9: EFC formula inputs and construction of empirical counterparts

FAFSA category	FAFSA line number and description	Construction	Data source	Note
I. Parents' income in 2012	7. Total Income	[1]	HSLS:09	If negative, set to 0.
II. Allowances against parents' income	8. U.S. Income tax paid if tax filers	[2]	HSLS:09, CPS	Assume everyone files taxes.
	9. State and other tax allowance	WT A1		WT A1 value for blank state.
	10. Parent 1 Social Security tax allowance	WT A2		WT A2 value at #7.
	11. Parent 2 Social Security tax allowance	WT A2		WT A2 value at #7.
	12. Income protection allowance	WT A3 [3]	HSLS:09	
	13. Employment expense allowance	[4a,4b]	HSLS:09	
	14. Total allowances	# 8 to # 13		
III. Available income	15. Available income (AI)	#7 to #14		May be negative.
IV. Parents' contribution from assets	20. Net worth	[5]	SCF	#s 16 to 19 are not imputed.
	21. Ed. saving and asset protection allowance	WT A5 [6]	HSLS:09	WT A5 value for blank parent age.
	22-24. Contribution from assets	[7]		If negative, set to 0.
V. Parents' contribution	25. Adjusted Available Income (AAI)	#15 + #24		
	26. Total parents' contribution from AAI	WT A6		Evaluate WT A6 using AAI.
	27. Number dep. in college in 2013-2014	[8]	HSLS:09	If not directly observed set to 1.
	28. Parents' contribution	$\frac{\#26}{\#27}$		If negative, set to 0.
VI. EFC		Equal to #28		Set contributions from student to 0.

[1] HSLS:09 variables: X1FAMINCOME,X2FAMINCOME

[2] Tax liability $y - \hat{\lambda} y^{1-\hat{\tau}}$, $\hat{\lambda}$ and $\hat{\tau}$ estimated using CPS and y from [1].

[3] HSLS:09 variables: X2HHNUMBER.

[4a] HSLS:09 variables: P2MARSTAT,X2PAR1EMP, X2PAR2EMP.

[4b] FAFSA formula: $\min\{0.35y, 3900\}$ where y is lowest parent earned income.

[5] Source: see appendix text.

[6] HSLS:09 variables: P2MARSTAT

[7] FAFSA formula: 12 percent of # 20 - # 21.

[8] HSLS:09 variables: P2INCLG2013,P2SIBSTARTCLG, and P2SIBCLGGRAD

Notes: Table A9 summarizes our imputation of the EFC by mapping2013 FAFSA formula inputs to empirical counterparts. The first column reports the FAFSA formula category; the second column provides the FAFSA line number and short description. The third column explains how the empirical counterpart is constructed; for explanations that are longer a table note number is provided in square brackets. The fourth column explains the data source for the empirical counterpart and the fifth column provides additional notes. Acronyms: WT refers to Worksheet Table, HSLS:09 refers to the High School Longitudinal Study of 2009, the CPS refers to the Current Population Survey Annual Socioeconomic Supplement, and SCF refers to the Survey of Consumer Finances.

Based on publications from the Federal Student Aid Office of the U.S. Department of Education and the National Association for Financial Aid Administrators, the Pell Grant is awarded first, and other sources of aid awarded subsequently, when financial aid administrators (FAAs) generate aid package offers for an applicant to their school (Program Communications Division, Federal Student Aid, 2013; NASFAA Monograph, 8th Edition, 2023; Office of Postsecondary Education, Federal Student Aid, 2013). In our imputation the amount of Pell grants that a high school graduated is eligible for is assigned using the difference between an upper bound and the imputed EFC (if negative this difference is set to 0). The upper bound is the minimum of the maximum Pell grant amount for that year, \$5,645, and the COA that we impute to the individual as described above.

Validating imputed postsecondary financial aid eligibility We validate the imputed Pell grant amount that individuals in the HSLS:09 are eligible for by comparing imputed Pell grants with realized Pell grant amounts for 2013-2014 BA enrollees who apply for aid in the HSLS:09 in the 2014 academic year. The realized aid values are pulled from student records submitted by postsecondary institutions to the HSLS:09 data collectors and are therefore likely to be accurate. Note that we focus on Pell grants because by comparison the contribution of other components (i.e., subsidized Stafford loans) to the total aid package depends on the discretion of the individual FAA who is designing the aid package. Table A10 presents moments of the distribution of imputed Pell grants, realized Pell grants, and the individual-level difference between the two in the population of 2013-2014 BA enrollees who submitted a FAFSA. Panel B further restricts the sample to those for whom imputed Pell grants are positive. The median difference in Panel A is 0 dollars, and the median difference in Panel B is -50 dollars.

Table A10: Imputed Pell grants compared to realized values

Panel A: 2013 BA enrollee FAFSA filers	Mean	p10	p25	p50	p75	p90
Realized	1,563	0	0	0	3,692	5,645
Imputed	1,452	0	0	0	2,796	5,645
Difference: Realized - Imputed	110.5	-1,749	0	0	0	2,996
Obs	2,266					
Panel B: ... and eligible for Pell grants	Mean	p10	p25	p50	p75	p90
Realized	3,517	0	0	4595	5,645	5,645
Imputed	4,822	1,964.6	5645	5645	5,645	5,645
Difference: Realized - Imputed	-1,305	-5,645	-2891	-50	0	881
Obs	577					

Notes: Table A10 compares realized and imputed values for Pell grants received by 2013 BA enrollees who filed a FAFSA (Panel A) and those who are additionally eligible for Pell grants according to the imputation (Panel B). Values are current dollars. Sample: bachelor's degree enrollees who submitted a FAFSA. Weights: PETS-SR longitudinal weights. Source: HSLS:09.

Although our imputation exhibits a high degree of accuracy, it is clearly not perfect. There are several reasons why this may be the case, most of which have an ambiguous effect on the accuracy of the Pell grant imputation. Sources of error that have an ambiguous effect on the accuracy of imputed Pell grant eligibility include the fact that, in the public HSLS:09 we observe family income as a discretized categorical variable, but have to assign a dollar value to each individual in order to proceed. We do this by using the midpoint of each category as the dollar value of income. Because the income bins represent wide intervals of income, this introduces error which is likely reflected in the imputed EFC and thus the imputed Pell grant eligibility. It is also true that parent income might fluctuate from the second follow-up to the FAFSA filing period. Similarly, we assume that the employment status of parents stays the same between the second follow-up and FAFSA filing. Other sources of error that have an ambiguous effect are our measure of room and board, used to assign cost of attendance to enrollees, which is a flat rate and not observed at the individual level. We also impute assets as a flat amount from the SCF. We also impute state tax liability which erases state-level variation in this contributor to the EFC. Sources of error that introduce an upward bias in imputed Pell eligibility include our assumption that the student contribution is 0. We also assume that siblings who are enrolled in college in the second follow-up are still enrolled when the respondent enrolls in college. If these assumptions are incorrect, the total EFC is higher and the Pell grant the student is eligible for is lower.

A.1.4 FAFSA nonfiling: reasons and money left on the table

Data description: possible reasons for not filing a FAFSA The possible reasons that survey respondents may give in the HSLS:09 for not filing a FAFSA are:

1. Because did not want to go into debt
2. Because can afford college without financial aid
3. Because thought ineligible or unqualified
4. Because did not know how
5. Because forms were too time-consuming or too much work
6. Because did not know could
7. Because teen does not plan to continue education

Respondents may respond yes/no for each reason to indicate that the reason applies; more than one reason may apply for the same survey respondent. If they respond that they thought they were ineligible or unqualified, further questions are asked: why did you think you were ineligible? Options provided to respondents are:

3. Thought ineligible or unqualified...
    (a) Because other family member didn't qualify
    (b) Because of concerns about credit score
    (c) Because income too high
    (d) Because grades/test scores too low
    (e) Because of part-time enrollment

In the main text, we group these responses into three categories: first, responses 1,3, and 6 as related to beliefs about eligibility; second, responses 4 and 5 as related to fixed costs; and, third, responses 2 and 7 as related to other reasons not related to possible frictions.

Subjective beliefs about Pell grant eligibility by family income tercile Table A11 reports, by family income tercile and postsecondary education outcome, the share of FAFSA nonfilers within each postsecondary outcome that did not file a FAFSA because they believed they were ineligible and the share that were mistaken about financial aid eligibility. This table uses the same definition as the main text of believing oneself ineligible for financial aid. Mistakes about eligibility are predominately in the bottom two income terciles of the income distribution among non-enrollees. Because those who actually enroll in postsecondary education tend to file FAFSAs quite often, especially for 4-year programs, the sample counts of FAFSA-nonfilers by income bin are quite small for enrollees.

Table A11: FAFSA non-filer because thought ineligible, didn't want debt, didn't know could, by income tercile

Income tercile	Variable	Did not enroll	l.t. 2-year	AA (2-year)	BA (4-year)
1	Share answering Yes	0.17	0.08	0.05	0.01
		(0.014)	(0.024)	(0.011)	(0.005)
		[692]	[125]	[405]	[578]
	... and eligible for aid	0.16	0.07	0.04	0.01
		(0.014)	(0.023)	(0.010)	(0.005)
		[692]	[125]	[405]	[578]
2	Share answering Yes	0.21	0.10	0.09	0.03
		(0.018)	(0.022)	(0.013)	(0.005)
		[494]	[185]	[484]	[1140]
	... and eligible for aid	0.04	0.02	0.03	0.01
		(0.009)	(0.011)	(0.007)	(0.002)
		[494]	[185]	[484]	[1140]
3	Nonfiler because believed ineligible	0.25	0.29	0.17	0.13
		(0.030)	(0.035)	(0.020)	(0.007)
		[209]	[166]	[367]	[2200]
	... and eligible for aid	0.00	0.07	0.01	0.00
		(0.000)	(0.020)	(0.006)	(0.001)
		[209]	[166]	[367]	[2200]
	Observations	1,395	476	1,256	3,918

Notes: Table A11 reports the share who belief they are ineligible for financial aid (using the definition of the main text) and the share who are mistaken, by income bin (rows) and postsecondary outcome in the fall of 2013 (columns). Moving down a column, shares are followed by standard errors in parentheses and within-group sample counts in square brackets. Weights: Second Follow-up student longitudinal weights. Source: HSLS:09.

Reasons for not submitting a FAFSA among high school graduates Table A12 begins by reporting the frequency of a more restrictive definition for not filing a FAFSA due to beliefs about eligibility for financial aid; specifically, in Panel A we show the frequency of responding either that they thought they were ineligible or they did not know they could. According to this definition, 7 percent of non-enrollees are mistaken about eligibility for financial aid. If not wanting to go into debt is considered by the reader as too loosely related to believing onself ineligible, this panel indicates that there is a sizable mass of mistaken beliefs remaining after dropping this from the definition. In Panel B, we tabulate the frequency of giving a reason for not filing a FAFSA related to either type of friction (beliefs or fixed costs of filing). The share of non-enrolles constrained by such frictions is 12 percent. In Panel C, we generate a flag that requires *only* answering the responses related to beliefs about eligibility from the main text; this sets the flag to zero for those who also affirmed that other reasons played a role. Panel D performs the same exercise but for fixed-cost responses. In both Panels, the share falls relative to more expansive definitions but remains positive and statistically significant.

Table A12: Reasons for not submitting a FAFSA among high school graduates

	Enrollment status in fall of 2013			
Description	**Did not enroll**	**l.t. 2-year**	**AA (2-year)**	**BA (4-year)**
Panel A: thought ineligible or did not know could				
Share answering Yes	0.14	0.13	0.08	0.07
	(0.009)	(0.015)	(0.007)	(0.004)
... and eligible for aid	0.07	0.05	0.03	0.00
	(0.007)	(0.010)	(0.005)	(0.001)
Panel B: any response except can afford and not planning to continue education				
Share answering Yes	0.23	0.17	0.10	0.08
	(0.011)	(0.017)	(0.008)	(0.004)
... and eligible for aid	0.12	0.06	0.04	0.01
	(0.009)	(0.011)	(0.005)	(0.001)
Panel C: only answering thought ineligible, did not know could, or did not want debt				
Share answering Yes	0.06	0.05	0.02	0.02
	(0.006)	(0.010)	(0.004)	(0.002)
... also eligible for Pell grant aid	0.04	0.02	0.01	0.00
	(0.005)	(0.007)	(0.003)	(0.001)
Panel D: only answering too hard or did not know how				
Share answering Yes	0.03	0.01	0.01	0.00
	(0.005)	(0.004)	(0.002)	(0.001)
... also eligible for Pell grant aid	0.02	0.01	0.01	0.00
	(0.003)	(0.003)	(0.002)	(0.000)
Obs	1,395	476	1,256	3,918

Notes: Table A12 reports the share of FAFSA nonfilers who report not filing for various reasons, by postescondary outcome in the fall of 2013. See options listed in Appendix A.1.4 for list of all possible choices. Weights: Second Follow-up student longitudinal weights. Source: HSLS:09.

Mistaken beliefs about eligibility: comparison of attributes and outcomes with those not misinformed Table A13 compares the composition of those who did not enroll by misinformation status about eligibility for financial aid. Those misinformed are slightly lower skill, more likely to be female, and tend to come from poorer families. They are less likely to have 2 resident parents, and more likely to have a sibling enrolled in college when they finish high school. The share of the misinformed who enroll in any postsecondary education by 2016 is less than 50 percent of the share among the not-misinformed who later enroll.

Binding fixed cost of filing a FAFSA: comparison of attributes and outcomes with those not constrained by fixed costs Table A14 compares attributes of FAFSA nonfilers who report being

Table A13: Summary statistics by misinformed status among 2013 non-enrollees

Variable	**Misinformed**	**Not misinformed**
High school GPA	2.3685	2.4049
Female	0.5356	0.4003
Family income	26,329	54,372
Below 185 % poverty line	0.8367	0.4831
2-parent family	0.6908	0.8522
Household size	4.1228	4.3124
Number siblings in PSE	0.4932	0.3753
Enrolled in any PSE 2016	0.0592	0.1349
Obs	135	1,260

Notes: Weights: Second follow-up student longitudinal weights. Source: HSLS:09.

affected by the fixed cost of filing with those who are not affected among 2013 non-enrollees. Here, being affected means you do not file a FAFSA and you were eligible for a positive Pell grant according to our imputation. Those affected by fixed costs have slightly higher GPAs, are more likely to be female, come from poorer families, are less likely to live in a single-parent family, have similar family sizes but more siblings in college, and are less likely to enroll in any kind of post-secondary education by 2016.

Table A14: Summary statistics by whether or not affected by fixed cost

Variable	**Affected by fixed cost**	**Not affected by fixed cost**
High school GPA	2.4329	2.3991
Female	0.5677	0.4038
Family income	29,449	52,996
Below 185 % poverty line	0.7777	0.5018
2-parent family	0.7969	0.8385
Household size	4.3129	4.2918
Number siblings in PSE	0.4401	0.3838
PSE enrollment in 2016	0.0446	0.1327
Obs	88	1,307

Notes: Weights: Second Follow-up student longitudinal weights. Source: HSLS:09.

Frequency of later applications for financial aid among 2014 AY non-filers Table A15 reports the frequency of FAFSA non-filers who did not enroll in the fall of 2013 (the 2014 AY) and who later are observed in post-secondary institutional records as having submitted a FAFSA in either the 2015 or 2016 AY. This rate is quite low.

Table A15: Share filing FAFSA in later years, among 2014 AY FAFSA non-filers who didn't enroll

Variable	**2015 AY**	**2016 AY**
Share	0.005	0.006
Obs	902	902

Notes: Weights: Second Follow-up student longitudinal weights. Source: HSLS:09.

A.1.5 Financial aid receipt among 2013 enrolles - extensive and intensive margin

Table A16 reports the share receiving aid, and the amount conditional on receipt, for Pell grants and subsidized Stafford loans among 2013 BA enrollees for whom we have a student record with nonmissing information about student aid receipt (which includes those who receive no aid).

Table A16: Financial aid among 2013 BA enrollees - Pell grants and subsidized Stafford loans

Pell grants		Subsidized Stafford SL	
Share receiving	**Ave amt. | receipt**	**Share receiving**	**Ave amt. | receipt**
0.332	4,413	0.471	3,221
Obs	2,558		

Notes: Weights: PETS-SR longitudinal weights. Source: HSLS:09.

A.1.6 Financial aid accounting

Table A17 reports the distribution of different categories of postsecondary financial aid across postsecondary education program types (Panel A) and within program types (Panel B). These distributions are computed using the HSLS:09 cleaned sample, and they are useful for motivating our model specification in the main text. In Panel A, it is evident that most of each type of aid is spent on BA programs. This is the type of postsecondary education program that we model in our framework. In Panel B, within BA programs, financial aid received by first year enrollees in the 2013-2014 academic year is similarly sourced from the three programs that we account for in our model specification of the main text: Pell grants and Stafford loans (subsidized or unsubsidized).

Table A17: Financial aid accounting: distribution of realized aid in the 2014 AY

		Postsecondary program type		
Distribution of aid	**Type of aid**	**l.t. 2-year**	**AA (2-year)**	**BA (4-year)**
Panel A: across PSE types	Pell grants	0.07	0.37	0.55
	Subsidized Stafford loans	0.04	0.13	0.82
	Unsubsidized Stafford loans	0.06	0.14	0.81
Panel B: within PSE types	Pell grants	0.51	0.65	0.31
	Subsidized Stafford loans	0.19	0.16	0.32
	Unsubsidized Stafford loans	0.30	0.19	0.36
	Observations	3,607		

Notes: Table A17 reports distributions of financial aid by aid type both across postsecondary program types (Panel A) and within program types (Panel B). Rows in Panel A sum to 1; Columns in Panel B sum to 1. Sample count reflect valid observations of financial aid outcomes among each postsecondary outcome. Weights: PETS-SR longitudinal weights. Source: HSLS:09.

A.1.7 Financial aid eligibility: estimating federal tax liability using CPS ASEC

We use the Current Population Survey Annual Socio-economic Supplement (CPS ASEC) to estimate federal income tax *liability* as a function of income, which is an input into the synthetic FAFSA formula used to impute EFC and Pell grant eligibility.

Federal tax liability: estimation approach and sample We are interested in estimating values for, $\hat{\tau}_y$ and $\hat{\lambda}_y$, the two parameters in the tax liability formula $y - \lambda_y (y)^{1-\tau_y}$ from Heathcote, Storesletten, and Violante (2017) that we impose on the HSLS:09 data as part of our imputation of means-tested financial aid eligibility. This functional form for the tax system implies that after tax income is equal to $\lambda_y (y)^{1-\tau_y}$, where y is pretax income. Equations (13) and (14) show the relationship after taking logs and the estimation equation, respectively. The estimation equation contains a normally distributed noise term, ϵ, as well.

$$\ln(\text{after-tax income}) = \ln(\lambda_y) + (1-\tau_y)\ln(\text{pretax income}) \tag{13}$$

$$\ln(\text{after-tax income}) = \text{Constant} + \beta \ln(\text{pretax income}) + \epsilon \tag{14}$$

To approximate the federal income tax liability of families in the HSLS:09 by estimating equation (14), we turn to the Current Population Survey's Annual Socio-economic supplement for 2013 (CPS ASEC). This survey is representative of the civilian noninstitutionalized population over the age of 16 (U.S. Bureau of Labor Statistics and U.S. Census Bureau, 2006). Observations of individuals are grouped into households in the survey; we also break down households into tax units using a method based on Lin (2022), in order to be consistent with the methodology used by the CPS to generate the tax variables provided by the survey to researchers. Specifically, we assign tax units within a household by first assigning individuals who are either the household head, their spouse, or their dependent child, foster child, or grandchild who is under the age of 18, not a tax filer, and who is unmarried at the time of the survey to tax unit 1. We then allocate remaining residents of the household to this original tax unit if they are the adult children of adults in that tax unit who are either under 18, or under 24 and enrolled in school, or receiving disability income, and who are not a tax filer. If there are additional individuals within the household who are not part of the original tax unit, we assign them to additional tax units by checking for married couples, who are grouped together in an additional tax unit. Finally, remaining individuals in the household who are not part of any tax unit at this point are assigned their own individual tax unit. After assigning individuals to tax units, following Lin (2022), we flag those who are not enrolled in school and who are either the household head, their spouse, or their cohabiting partner, or who are the child, foster child, or grandchild of the household head and who are over the age of 18. We consider this the sample of adults. We then flag tax units where there is at least one adult who is between the ages of 20 and 64. This is our estimation sample for the estimation of federal tax liability as a function of income.

The CPS reports adjusted gross income for tax units assigned with the CPS methodology described in Lin (2022); we use the log of this value as the pretax income variable for the tax progressivity estimation (the independent variable). We subtract the federal tax liability after all credits of the tax unit from the adjusted gross income of the tax unit and take the log, in order to construct the after-tax variable for the tax progressivity estimation (the dependent variable). We then run an ordinary least squares regression of the independent on the dependent variable using ASEC household weights, on the estimation sample that we construct as described above. Table A18 reports summary statistics for our estimation sample.

Table A18: CPS sample summary statistics

	Mean	Median
Adjusted gross income	68945	48000
Federal tax liabiliaty after credits	8125	3503
Age	43	43
Observations	61,305	

Notes: Table A18 reports estimation sample summary statistics for the tax progressivity estimation. Sample: households with household head or spouse aged between 20 and 65 and at least one adult who is not enrolled in school and who is either the household head, spouse, adult child or grandchild, or cohabiting partner of the household head. Weights: ASEC household weights. Source: CPS ASEC.

Federal tax liability: estimation results Table A19 reports our results for our estimation imposing equation (14) on the data. The progressivity parameter and average tax parameter estimates, $\hat{\tau}_y$ and $\hat{\lambda}_y$, are shown at the bottom of the table. To estimate federal tax liability, we evaluate the expression for after tax income using observed income in the HSLS:09 and the parameters reported here. We then take the difference between pretax and after-tax income to find federal tax liability before transfers.

A.2 The PSID and NLSY97

We use the Panel Study of Income Dynamics (PSID) and the 1997 National Longitudinal Survey of Youth (NLSY97) to estimate the deterministic and stochastic components of the life cycle earnings process. Our estimation approach follows Moschini, Raveendranathan, and Xu (2023) with the modification that skill is measured in the NLSY97 using high school GPA rather than the AFQT. The PSID estimation sample is identical to that paper; the NLSY97 estimation sample cleaning procedure is identical but requires that we observe high school GPA, which slightly

Table A19: CPS federal tax progressivity estimation results

Variable	Coefficient
Ln(pretax income)	0.919 (0.000367)
Constant	0.816 (0.00393)
Observations	61,305
Results	
$\hat{\tau}_y$	0.081
$\hat{\lambda}_y$	2.261

Notes: Table A19 reports results for the tax progressivity estimation. Standard errors in parentheses. Sample: see text. Weights: ASEC household weights. Source: CPS ASEC.

affects estimation sample counts.

A.2.1 Life cycle earnings process functional forms and estimation results

The deterministic component $\epsilon_{j,e,s}$ depends on the consumer's age, j, their education, e, and their skill endowment, s:

$$\epsilon_{j,e,s} = \exp\left(\beta_{e,1}^{A} j + \beta_{e,2}^{A} j^{2} + \beta_{e,3}^{A} j^{3} + \beta_{e,s}^{s}\right)$$

The stochastic component is an AR(1) process where the persistence parameter ρ_e depends on the consumer's educational attainment, as does the variance σ_e^2 of the Normal distribution from which the error term is drawn:

$$\begin{aligned} \eta^{'} &= \rho_e \eta + \nu_e \\ \nu_e &\sim \mathbb{N}\left(0, \sigma_e^2\right) \end{aligned}$$

For each education group e, we use the Panel Study of Income Dynamics (PSID) to estimate how logged real wages are related to a third-order polynomial of age within each education group. This identifies $\beta_{e,1}^{A}$, $\beta_{e,2}^{A}$, and $\beta_{e,3}^{A}$. The longer panel dimension of the PSID makes it better suited to this task than the NLSY97. We then clean logged hourly real wages in the NLSY97 of age effects with the PSID estimation results. We regress the resulting age-free log hourly real wages on indicators for skill terciles; the estimated coefficients on skill tercile indicators are the factor loadings on skill s, $\beta_{e,s}^{s}$. The fact that high school GPA is observed in the NLSY97 but not the PSID makes the NLSY97 suited to this task. The residuals from the NLSY97 regression are then used to jointly estimate ρ_e and σ_e^2. Point estimates are reported in Table A20.

Table A20: Earnings process estimation results

Parameter	Description	Value given education e	
		$e = \ell$	$e = h$
$\beta^A_{e,1}$	Age third-order polynomial	0.106	0.186
$\beta^A_{e,2}$		-0.00176	-0.00319
$\beta^A_{e,3}$		0.00000897	0.0000174
$\beta^s_{e,1}$	Skill endowment tercile shifter	-0.0396	-0.201
$\beta^s_{e,2}$		-0.0329	-0.122
ρ_e	Persistence AR(1)	0.851214	0.886269
σ^2_e	Variance AR(1)	0.082873	0.072023

Notes: Table A20 summarizes the results from the earnings process estimation. Sources: PSID and NLSY97.

A.2.2 College wage premium by skill tercile

Table A21 reports the college wage premium in the NLSY97.

Table A21: Bachelor's degree wage premium by skill tercile: ratio of median wages

	High school		Bachelor's degree		
Skill	Wage	Obs	Wage	Obs	Wage premium
1	14.12	6854	18.63	870	1.32
2	14.67	5546	20.85	2402	1.42
3	14.92	2687	23.41	5350	1.57

Notes: Table A21 tabulates the median wage within each skill tercile by education attainment status for those not currently enrolled in post-secondary education; the last column is the ratio of median wages in the two educational attainment categories, the college wage premium. Observation counts are at the individual-year level. Source: NLSY97.

A.2.3 Inter-vivos transfers

To estimate the average ratio of transfers to income in the NLSY97, we use a subset of the earnings process estimation sample. To estimate average inter vivos transfers from parents to their college-aged children in the NLSY97, we use the earnings process estimation sample with four modifications involving requirements on respondent age, education status, independence status, and whether the observation has been assigned a family/parent income tercile. First, we allow individuals to be enrolled in an education program in a given year; second, we restrict attention to individuals classified as independent by the NLSY97 in a given year; third, we keep individuals between the ages of 18 and 23 during the years from 1997 to 2003; and, fourth, we require that we observe family income tercile. This leaves 8,291 individual-year observations (3,063 individuals).

To account for an implicit transfer from parents to their children who cohabit with them and do not pay rent, we flag those cohabiting with their parents and paying no monthly rent, then impute the average monthly rent paid by sample members with the same family income tercile, college enrollment status, and observation year who are not cohabiting. Next, we transform monthly rent to yearly rent, and add it to yearly net income received from parents (if both parents are present) or from both the mother and father (if both parents are not present). We also add any yearly allowances received. The resulting quantity is the yearly nominal transfers from parents to their child. Within each year, we then multiply the quantity by 6 and divide by nominal GDP per capita in that year (for those over 18) to find a unitless implied ratio of transfers received to per capita income for each individual while they are young adults of college age. To compute GDP per capita for those 18 and over in these years, we use we normalize by GDP per capita for those 18 and over, which is computed by combining information on GDP from 1997-2003 from BEA (2022, T1.1.5) and population levels from the US Census Bureau found in Census Bureau of the United States (2000, 2010). We then average this ratio across invididuals and years to find the value reported in the first row of Table A22. The average real values of the components of transfers are also reported. To convert current dollars to 2012 US dollars, we use the Consumer Price Index (CPI).

Table A22: Inter vivos transfers

Variable	**Mean**
Transfer ratio	0.579
Transfers	6,281
Transfers not allowance	721
Allowance	188
Imputed rent	6,227
Obs (invidual-year)	8,291
Obs (inviduals)	3,063

Notes: Table A22 reports average transfers for the sample used to estimate inter vivos transfers. Sample: independents between 18 and 23 observed during 1997-2003. Units for transfer amounts: year 2012 USD. Data are at the individual-year level. Source: NLSY97.

A.3 Income tax progressivity τ_p: estimation data from the CBO and results

In order to estimate the progressivity of the income tax and transfer system, we use data unerlying figures from the CBO reports on the distribution of household income (U.S. Congressional Budget Office, 2018a,b), following the robustness exercise of Heathcote, Storesletten, and Violante (2017). Note that, although we parameterize to the period 2013-2015, in this estimation we omit 2013 because the CBO report for that year does not include the components necessary to implement the estimation procedure described below.

Qualitatively, the income tax progressivity parameter, τ_p, represents the progressivity of the federal tax system including transfers from Temporary Aid for Needy Families (TANF), SNAP (Supplemental Nutrition Assistance Program), and SSI (Supplemental Security Income). This is in contrast to the estimation in Appendix A.1.7 which only establishes tax liability for the purposes of imputing the EFC and Pell grant eligibility.

Table A23 reports the baseline federal tax rate, as well as the transfer rates from Temporary Assistance to Needy Families (TANF), Supplemental Nutrition Assistance Program (SNAP), and Supplemental Security Income (SSI) shown in columns (1), (2), (3), and (4), respectively. We compute the empirical equivalent of the net tax rate for our model as the federal tax rate (which includes refundable credits as reported in column 1) minus the transfer rates from TANF, SNAP, and SSI and report this net tax rate in column (5). Average pretax income in column (6) is logged in column (7); logged after-tax income reported in column (8), where after-tax income is computed by taking the log of the net tax rate in column (5) applied to the pretax income of column (6).

Table A23: CBO data by year

Year	Percentiles		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Min	Max	Fed. tax	TANF	SNAP	SSI	Net tax	Ave. Y	$\log(Y)$	$\log(Y_{AT})$
2014	99	100	33.6				33.6	1.776	0.249	0.071
	96	99	26.8				26.8	0.251	-0.466	-0.601
	91	95	23.4				23.4	0.158	-0.685	-0.801
	81	90	21.2				21.2	0.119	-0.820	-0.924
	60	80	17.8				17.8	0.104	-0.981	-1.066
	40	60	14.0				14.0	0.059	-1.163	-1.229
	20	40	9.1	1.6	1.6	1.1	4.8	0.040	-1.376	-1.397
	0	20	1.9	6.6	9.7	7.0	-21.4	0.023	-1.717	-1.632
2015	99	100	33.3				31.3	1.237	0.268	0.092
	96	99	26.7				26.7	0.261	-0.449	-0.583
	91	95	23.6				23.6	0.163	-0.670	-0.786
	81	90	21.3				21.3	0.124	-0.804	-0.908
	60	80	17.9				17.9	0.089	-0.967	-1.052
	40	60	14.0	0.5			13.5	0.061	-1.149	-1.212
	20	40	9.2	2.0	1.1	0.9	5.3	0.042	-1.357	-1.380
	0	20	1.5	9.7	8.1	5.9	-20.7	0.024	-1.699	-1.617

Notes: Table A23 reports the components for the estimation of the income tax progressivity parameter τ_y. Data is from 2014 and 2015, and dollar values in column (6) are in millions of current USD. After-tax income is defined as $Y_{AT} \equiv \left(1 - \frac{\text{Net tax}}{100}\right) Y$, where the net tax rate is defined as $(5) \equiv (1) - (2) - (3) - (4)$.

Table A24 reports the estimation results using CBO data presented in Table A23.

Table A24: Income tax progressivity estimation results by year and overall

Coefficient	**2014**	**2015**
β_1	0.822 (0.0314)	0.824 (0.0311)
β_0	-0.251 (0.0332)	-0.246 (0.0323)
Implied $\hat{\tau}_{p,t}$	0.178	0.176
Average 2014-2015 $\hat{\tau}_p$	0.177	

Notes: Table A24 reports estimation results. Standard errors in parentheses.

A.4 Consumption tax τ_c: estimation data from the OECD and results

In order to estimate the consumption tax rate, τ_c, we apply equation (5) from Mendoza, Razin, and Tesar (1994) to updated data for our parameterization time period. This equation is:

$$\tau_{c,t} = 100 \times \frac{5110_t + 5121_t}{C_t + G_t - GW_t - 5110_t - 5121_t} \quad (15)$$

Specifically, we use values for the United States from three data series (OECD, 2024c,b,a) to populate the 2013, 2014, and 2015 entries of Panels A, B, and C in Table A25.

Table A25: OECD data by year

Variable	Description	2013	2014	2015	Source
Panel A: Total tax revenue (all levels of government)					
5110	General taxes on goods and services	343,853	361,685	374,173	OECD (2024c)
5121	Excises	154,390	155,976	156,902	
Panel B: Final consumption expenditure					
C	Private	11,040,849	11,521,194	11,933,651	OECD (2024b)
G	Government	2,530,745	2,562,276	2,603,988	OECD (2024a)
Panel C: Compensation of employees by source					
GW	Paid by producers of gov't services	1,665,524	1,706,888	1,758,064	OECD (2024a)

Notes: Table A25 reports OECD data used in the consumption tax rate estimation method of Mendoza, Razin, and Tesar (1994). Dollar values are in millions of current USD for that year, rounded to the nearest dollar.

Results are presented in Table A26. The average rate across the 2013-2015 time period is 0.044, which is the number we assign to τ_c.

Table A26: Consumption tax rate estimation results by year and overall

Variable	Description	2013	2014	2015
$\hat{\tau}_{c,t}$	Annual rate (share)	0.044	0.044	0.043
$\hat{\tau}_c$	Average rate 2013-2015 (share)	0.044		

B Model Appendix

B.1 Value functions

Consumers begin student loan payments the year after college graduation age, regardless of whether or not they complete college. The idiosyncratic state of a consumer while $j > 4$ and $j \neq j_f + j_a$ is given by the tuple (j, e, s, η, a). The consumer's value function is given by

$$V(j, e, s, \eta, a) = \max_{d_f \in \{0,1\}} (1 - d_f)V^R(j, e, s, \eta, a) + d_f V^D(j, e, s, \eta, a) \tag{16}$$

where d_f denotes the student loan delinquency decision. The terms $V^R(\cdot)$ and $V^D(\cdot)$ denote the value of repayment and the value of delinquency, respectively. The value of repayment for $j > 4$ and $j \neq j_f + j_a$ is given by

$$\begin{aligned}
&V^R(j, e, s, \eta, a) = \max_{c \geq 0, a'} U(c, j, e) + \beta\psi_j E_{\eta'|e,\eta} V(j + 1, e, s, \eta', a', x') \qquad (17)\\
&s.t.\\
&(1 + \tau_c)c + a' = y_{j,e,s,\eta,a} + a + \mathbb{I}_{\{a<0\}} r_{SL} a + Tr_j - T(y_{j,e,s,\eta,a})\\
&a' \geq \min[(1 + r_{SL})a + \rho_R(j, a), 0]
\end{aligned}$$

If the consumer has outstanding federal loans, then $a' \geq (1 + r_{SL})a + \rho_R(j, a)$.

Alternatively, these consumers can choose delinquency. Their value function for $j > 4$ and $j \neq j_f + j_a$ is given by

$$\begin{aligned}
&V^D(j, e, s, \eta, a) = U(c, j, e) - \xi_D + \beta\psi_j E_{\eta'|e,\eta} V(j + 1, e, s, \eta', a') \qquad (18)\\
&s.t.\\
&(1 + \tau_c)c = y_{j,e,s,\eta,a} + Tr_j - T(y_{j,e,s,\eta,a}) - \rho_D(j, a, y_{j,e,s,\eta,a})\\
&a' = (1 + r_{SL})a + \rho_D(j, a, y_{j,e,s,\eta,a}) - \phi_D[\rho_R(j, a) - \rho_D(j, a, y_{j,e,s,\eta,a})]
\end{aligned}$$

where ξ_D is the stigma cost of choosing delinquency. In the case of non-repayment, consumers do not make a consumption-savings decision. Instead, they have their wage garnished to make a partial payment of $\rho_D(j, a, y_{j,e,s,\eta,a})$. Therefore, they consume whatever remains from their disposable income, plus accidental bequests, after making the partial payment. The parameter ϕ_D is the fraction of missed payment (difference between full payment and partial payment) that is charged as a collection fee. The outstanding principal plus interest is then augmented by the missed payment plus the collection fee (net of any partial payment).

When $j = j_f + j_a$, in addition to the choices described above, the parent chooses an inter vivos transfer to their child, who will become an independent agent. At the start of age $j_f + j_a$, the parent draws their child's skill type and the family's subjective belief and then chooses whether or not to be delinquent on any student debt payments; the expected family contribution of the child is observed given parental income. The value function before the draw of child skill type and subjective belief is given by

$$V(j,e,s,\eta,a) = \sum_{s_c} \pi_{s_c}(s_c) \sum_b \pi_b(b) \, [\max_{d_f \in \{0,1\}} (1-d_f) V^R(j,e,s,\eta,a,s_c,b) \tag{19}$$
$$+ d_f V^D(j,e,s,\eta,a,s_c,b)]$$

where $\pi_{s_c}(s_c)$ is the probability over child skill and $\pi_b(b)$ is the probability over the subjective belief about need-based financial aid eligibility. The value of repayment for $j = j_f + j_a$ is given by

$$V^R(j,e,s,\eta,a,s_c,b) = \max_{c \geq 0, a', a_c} U(c,j,e) + \beta \psi_j E_{\eta'|e,\eta} V(j+1,e,s,\eta',a') + \tag{20}$$
$$\beta_c E_{\eta'|\ell} W(s_c, \eta', a_c, EFC(y_{j,e,s,\eta,a}), b)$$
$$s.t.$$
$$(1+\tau_c)c + a' + a_c = y_{j,e,s,\eta,a} + a + r_{SL} a \mathbb{I}_{a<0} + Tr_j - T(y_{j,e,s,\eta,a})$$
$$a' \geq \min[(1+r_{SL})a + \rho_R(j,a), 0]$$
$$a_c \geq 0$$

where $EFC(y)$ is the expected family contribution formula defined in equation (3), a_c is the inter vivos transfer to the child, $W(\cdot)$ is the child's value function, and β_c disciplines the intensity of parental altruism toward the child. Because the parent uses $W(\cdot)$ for their child's lifetime utility, the parent also has the same subjective belief as the child about need-based financial aid. The child's AR(1) productivity η' is drawn from the stationary distribution for a consumer without a college degree.

When $j = j_f + j_a$ and the consumer chooses delinquency, we assume those consumers cannot make an inter vivos transfer to their child for simplicity. Therefore, the value functions for delinquency are the same as in equation (18), with the difference that the parent has a term reflecting altruistic utility toward their child, represented by the addition of $\beta_c E_{\eta'|\ell} W(s_c, \eta', a_c = 0, EFC(y), b)$ to the objective function.

B.2 Definition of equilibrium

To define the equilibrium, we must first discuss notation and define the function for Social Security transfers. Let $\overrightarrow{\omega}$ denote the idiosyncratic state of a consumer. This state depends on age and enrollment status in the following way:

$$\overrightarrow{\omega} = \begin{cases} (s, \eta, a, f, b) & \text{for 18-year-olds, before making the college entrance decision} \\ (j, h, s, \eta, a, f, b) & \text{for consumers in college} \\ (j, e, s, \eta, a) & \text{for consumers not enrolled, dropouts, or graduates, if } j \neq j_f + j_a \\ (j, e, s, \eta, a, s_c, b) & \text{if } j = j_f + j_a \end{cases} \tag{21}$$

Furthermore, let $d_{d,t}(\overrightarrow{\omega})$ denote the dropout decisions that solve the endogenous discrete dropout problems in the continuation values of equation (12).

Social Security transfer function: Social Security transfers replace a fraction χ of the average labor earnings for the 30 years before retirement conditional on education and skill plus the average unconditional labor earnings for the 30 years before retirement, divided by two. The transfer function is given by

$$ss_{e,s} = \frac{\chi}{2}\left[\frac{\int w_e \eta \epsilon_{j,e,s} \Omega_t d(\overrightarrow{\omega}|18 \leq j < j_r, e, s)}{\int \Omega_t d(\overrightarrow{\omega}|18 \leq j < j_r, e, s)} + \frac{\int w_e \eta \epsilon_{j,e,s} \Omega_t d(\overrightarrow{\omega}|18 \leq j < j_r)}{\int \Omega_t d(\overrightarrow{\omega}|18 \leq j < j_r)}\right] \tag{22}$$

Definition Given an initial level of capital stock K_0 and an initial distribution over idiosyncratic states $\Omega_0(\overrightarrow{\omega})$, a competitive equilibrium consists sequences of household value functions $\{W_t(\overrightarrow{\omega}), V_t(\overrightarrow{\omega}), V_t^R(\overrightarrow{\omega}), V_t^D(\overrightarrow{\omega})\}$, household college entrance and dropout policy functions $\{d_{e,t}(\overrightarrow{\omega}), d_{d,t}(\overrightarrow{\omega})\}$, household consumption and next period asset policy functions $\{c_t(\overrightarrow{\omega}), a'_t(\overrightarrow{\omega})\}$, household delinquency policy functions $\{d_{f,t}(\overrightarrow{\omega})\}$, household inter vivos transfer policy function $\{a_{c,t}(\overrightarrow{\omega})\}$, production plans $\{Y_t, K_t, L_t, L_{\ell,t}, L_{h,t}\}$, tax policies $\{\gamma_t\}$, prices $\{r_t, w_{\ell,t}, w_{h,t}$, Social Security transfers $\{ss_{t,e,s}\}$, accidental bequests $\{Tr_{t,j}\}$, and measures $\{\Omega_t(\overrightarrow{\omega})\}$ such that:

(i) Given prices, transfers, and policies, the value functions and household policy functions solve the consumer problems in equations (10)-(12) and (16)-(20);
(ii) The saving interest rate and wage rates satisfy equations firm first order conditions;
(iii) Social Security transfers satisfy equation (22);
(iv) Accidental bequests are transferred to households between ages 50 and 60 $(33 \leq j \leq 43)$ after deducting expenditure on private education subsidies[10]

$$Tr_{t+1,j} = \frac{\int (1-\psi_j) a'_t(\overrightarrow{\omega}) \Omega_t d(\overrightarrow{\omega}) - \kappa \int \theta^{pr}(s) \mathbb{I}_{e=h \text{ and } j \in \{1,2,3,4\}} \Omega_{t+1} d(\overrightarrow{\omega})}{\sum_{j=33}^{43} N_{t+1,j}} \tag{23}$$

[10] In our baseline calibration and in all of the counterfactual exercises, accidental bequests are always positive because the assets of those who die exceed the expenditure on private subsidies to education costs. If they did not exceed private subsidies, then bequests would be negative, which is equivalent to a lump-sum tax.

where $N_{t,j}$ denotes the mass of population of age j at time t;
(v) Government budget constraint balances as follows, by adjusting γ:

$$\int [\tau_c c_t(\overrightarrow{\omega}) + T(y_{t,j,e,s,\eta,a})]\Omega_t d(\overrightarrow{\omega}) = G_t + E_t + D_t + SS_t \tag{24}$$

where G_t, E_t, D_t, and SS_t are government consumption, total public education subsidy, federal student loan program expenditure, and Social Security expenditure;
(vi) Labor, capital, and goods markets clear in every period t; and
(vii) $\Omega_{t+1} = \Pi_t\Big(\Omega_t\Big)$, where Π_t is the law of motion that is consistent with consumer household policy functions and the exogenous processes for population, labor productivities, skill, subjective beliefs about need-based aid, and the probabilities of being allowed to continue college for each skill endowment bin.