Optimal Access to Federal Student Loans in the United States*

Emily G. Moschini[†] Gajendran Raveendranathan[‡] Ming Xu [§]

[PRELIMINARY]

This version: January 5, 2022

Abstract

Student loans, now the largest form of consumer credit after mortgages in the United States, have the highest delinquency rates among the major forms of consumer credit. Furthermore, consumers that never completed the program the loan was taken out to pay for (dropouts) account for more than half of delinquencies. Are students borrowing too much? We examine the optimal federal student loan limit in a life cycle model with over-optimism about college graduation likelihood. We discipline the model's positive relationship between student skill (high school GPA) and college graduation likelihood using panel microdata on high school students. In the calibrated model, over-optimism allows us to match shares of borrowing and delinquencies by dropouts. We find that, for the average student, expanding the student loan limit to fully pay for college is optimal. However, during the transition to this expanded limit, low-skill students experience welfare losses because they over-enroll in college.

JEL classification numbers: I22, I26, D83
Keywords: Post-secondary education, Student loans, Over-optimism, Student loan limits.

*We are grateful to Sergio Salgado for his help with the Survey of Consumer Finances, as well as Mons Chan, Eugenia Gonzalez, Hyunju Lee, Sergio Ocampo, Radek Paluzinski, and Thomas Phelan for their comments.

†College of William & Mary. E-mail: egmoschini@wm.edu. Website: https://sites.google.com/view/emilymoschini/.
‡McMaster University. E-mail: raveeg1@mcmaster.ca. Website: https://sites.google.com/site/gajendranraveendranathan/home.
§Queen's University. E-mail: ming.xu@queensu.ca. Website: https://www.mingecon.com/.

1 Introduction

Student loans are now the largest form of consumer credit after mortgages in the United States, with total outstanding balances amounting to 7.6 percent of GDP.[1] The payments imposed by this debt can be challenging for debtors to finance: in 2019, 19.4 percent of households with student loans were not making payments (i.e., "delinquent").[2] This situation has led some to ask whether students are being allowed or even encouraged to borrow too much (Athreya, Herrington, Ionescu, and Neelakantan (2021)), and even prompted calls for student loan forgiveness (Schumer and Warren (2020) and Friends of Bernie Sanders (2021)). Indeed, although the earnings premium from a college degree might be enough to make timely repayment possible, college completion is by no means guaranteed. In fact, households that did not complete the education the loan was taken out to pay for (henceforth dropouts), account for 18.35 percent of households with student loans and hold 10.79 percent of outstanding balances. These households in particular tend to struggle with student loan repayment: in 2019, more than 50 percent of delinquent households were dropouts.

In light of these facts, we study the optimal student loan limit. Under current US policy, the federal student loan limit for dependent students is sufficient to pay for 1.74 years of college.[3][4] Will expanding student loan limits increase or decrease well-being? We pose this question within an environment where college non-completion is partly explained by over-optimism at the point of enrollment, a modeling choice supported by evidence from the literature on the college dropout decision. In particular, in a panel study of students in Berea College, Stinebrickner and Stinebrickner (2012) find that the expected graduation rate was 86 percent, whereas the actual graduation rate was 65 percent.[5]

We use our model to analyze transition paths after changes in the student loan limit. Our main finding is that, for the average student, the optimal access to student loans is an expansion in borrowing limits up to the full cost of college. However, in the initial periods of the transition, some consumers with low high school GPA will be hurt by the expansion in limits, because for them over-optimism plays a large role in the enrollment decision.

Our main theoretical contribution is to incorporate over-optimism about graduation likelihood into a theory of college choice. We build on the general equilibrium life cycle framework of Krueger and Ludwig (2016) by augmenting their model with student loan repayment decisions (Luo and Mongey, 2019), endogenous and exogenous dropouts (Chatterjee and Ionescu, 2012), and over-optimism about graduation likelihood (Stinebrickner and Stinebrickner, 2012). Stinebrickner and Stinebrickner (2012) argue that especially students with low high school GPA are more over-optimistic about their likelihood of college graduation. We

[1]Source: Federal Reserve Board G19, Bureau of Economic Analysis, authors' calculations.

[2]Source: 2019 Survey of Consumer Finances, authors' calculations. See Appendix A.2, Table 9 for definition of households not making payments.

[3]Here, college cost is average net tuition and fees plus room and board, where net tuition and fees refers to tuition and fees after deducting all grants and subsidies (NCES (2019)).

[4]Most undergraduate students are dependents. This can be seen in NCES (2020), which indicates that in 2015-2016 73.9 percent of full-time, full-year postsecondary students were dependents, and that among those students in this group who received loans as financial aid, 74.2 percent were dependents.

[5]These findings are based on the Berea Panel Study, which is an on going survey that collects data on cohorts that entered Berea College in 2001 and 2002.

discipline the model's positive relationship between student skill (high school GPA), college enrollment, and college graduation likelihood using the High School Longitudinal Study of 2009, a nationally representative sample of 9th graders in the United States.

In our model, college is a four-year program. The benefit of college is higher expected earnings upon graduation, and a higher expected skill for one's future children. The costs of college are tuition, lower earnings while in college, and an effort cost. When making the college entrance decision, we assume that students are over-optimistic about their probability of college graduation. This over-optimism is captured by a belief about the minimum probability of graduation. Students will use the maximum of the minimum probability of graduation and the true probability of graduation when making their college entrance decision. Hence, the students will over-estimate the expected benefits of attending college. To pay for college, students may use inter vivos transfers from their parents, subsidies from the government and private beneficiaries, earnings from part time work, and borrow from the federal student loan program. In the model, the federal student loan program is characterized by a student loan limit, interest rate, and late payment penalties, similar to the current US system.

After calibrating the model to the US economy, we validate it against untargeted moments. First, we show that our model does reasonably well in explaining the extensive and intensive margins of student loan utilization, by showing that the model accounts for the distribution of student loan balances over high school GPA among fourth year students in college. Second, we show that our model predicts the shares of student loans and delinquent households accounted for by dropouts. We then use the model to quantify the importance of over-optimism in accounting for student loans and delinquencies among dropouts. Comparing our baseline model to one without over-optimism, we show that without over-optimism, the shares of student loans and delinquencies accounted for dropouts decreases by roughly 20 and 35 percent, respectively.

Having highlighted the quantitative significance of over-optimism in our model, we solve for the optimal access to federal student loans. In particular, we vary the loan limit common to all agents, and compute the transition paths that result. We then measure the welfare gains and losses for 18-year-olds indexed by their skill (high school GPA). When computing welfare, we assume that the government is paternalistic. That is, the government internalizes that the decision rule for college enrollment incorporates consumer over-optimism, but uses the correct payoff of these choices when it computes welfare. Whether a student experiences gains from a student loan limit expansion depends on their skill, as well as on whether they are born in the initial period of the transition or in the future after the transition is almost complete. Skill matters for the optimal borrowing limit because it determines the likelihood of over-enrollment and graduation at a given loan limit. The time period of the transition matters because of general equilibrium effects. In the long run, an expansion in limits will increase college enrollment, leading to more graduates. This results in a larger tax base, and hence, a lower income tax rate, which benefits all consumers.

Over-optimism causes over-enrollment, and thus welfare losses in the transition, in the middle of the skill distribution. Students with higher skill benefit from an expansion in borrowing limits because their true graduation probabilities are high enough for over-optimism not to affect their enrollment choices. Students with the lowest skill are not as likely to enroll in college regardless of what the loan limit is, while at the same time they benefit from lower income taxes that result from a significant expansion in limits. This is especially true in the long run. However, for students with low (but not the lowest) skill, over-optimism leads over-enrollment to increase. In the short run, gains from lower income taxes do not compensate for this, which leads to welfare losses for this group. Despite this heterogeneity in welfare gains, the average 18-year-old benefits from an expansion in loan limits.

Related literature Our paper complements several studies of the interaction between borrowing constraints and the college enrollment decision. The sizable college wage premium has generated interest across several economic fields in policies that can affect the enrollment choice. Usually, the literature uses rational-consumer structural models to study the general equilibrium effects of loan programs (Lochner and Monge-Naranjo (2011), Krueger and Ludwig (2016), Caucutt and Lochner (2020)), or microeconomic approaches in experimental or quasi-experimental settings to study interventions which may address frictions in the program uptake process (Dynarski, Libassi, Michelmore, and Owen (2021), Hoxby and Turner (2015), Stinebrickner and Stinebrickner (2012)). Our paper bridges these two approaches, by incorporating over-optimism into a general equilibrium policy analysis. The closest structural paper to ours is Krueger and Ludwig (2016), who analyze optimal policies for taxes and public education subsidies. The findings of Stinebrickner and Stinebrickner (2012) provide impetus for our main information friction (over-confidence about graduation among those of low skill).

We contribute the structural literature on student loans, which includes Ionescu (2009), Ionescu (2011), Chatterjee and Ionescu (2012), Ionescu and Simpson (2016), and Luo and Mongey (2019). First, we bridge this literature with the empirical literature on over-borrowing by incorporating students who are over-optimistic about graduation likelihood. Second, we focus on a policy related to access to student loans, whereas the policy analyses in these papers focus on student loan repayment policies. Third, to our knowledge we are the first in this structural literature to leverage the High School Longitudinal Study of 2009 to guide modeling choices related to a college education and student loans.

We also contribute to consumer credit literature on over-borrowing. Recently, several policies have been implemented in the market for consumer credit cards that are aimed at reducing access. For example, the Credit CARD Act of 2009 imposed several restrictions in the United States. In addition, Canada has recently imposed restrictions on borrowing. Some papers have analyzed over borrowing in the market for consumer credit cards. For example, Exler, Livshits, MacGee, and Tertilt (2021) analyze implications of policies aimed correcting for over-borrowing in the unsecured credit market due to over-optimism about earnings. Nakajima (2012) studies the impact of increased access to unsecured credit for consumers with hyperbolic discounting preferences. Nakajima (2017) studies how bankruptcy policy reforms in the unsecured credit market impact consumers with hyperbolic discounting preferences. In contrast to this literature, we focus on over-borrowing of federal student loans rather than in the consumer credit card market.

This paper proceeds as follows. Section 2 overviews our main empirical findings. Section 3 lays out the model, section 4 the calibration, and section 5 the main properties of the model's equilibrium, the policy experiment, and the results. Section 6 concludes.

2 Data

The two main data sets we draw on are the 2019 Survey of Consumer Finances (SCF) and the High School Longitudinal Survey of 2009 (HSLS). The HSLS follows a representative sample of high schoolers in the United States in 2009 through high school graduation, college enrollment (or not) and records persistence. However, the panel stops when high school graduates are at most in their fourth year of college. Because the panel dimension is limited, and we cannot observe repayment outcomes for student loans after college graduation, we supplement the information from the HSLS with the SCF. This second dataset allows us to observe how loan repayment delinquency and the distribution of student loan balances are associated with college completion in the US population.

Below we briefly describe each of these data sets and highlight our main findings from each data set. In particular, we use the SCF to demonstrate that student debt is sizeable, that the population of student debtors is quantitatively significant, that a sizeable fraction of student debtors did not complete their education, and that these non-completers are overrepresented among student debtors who are delinquent on their loan payments. We use the HSLS to show that student skill quintiles are correlated with college enrollment and college persistence by computing unconditional means for these outcomes within each quintile. By regressing student persistence on other attributes of the student (family income, parent education, student loan uptake, hours worked) as well as student skill, we show that even after controlling for other attributes of the student which could be associated with persistence in college, student skill remains significantly predictive of students remaining in college and not dropping out.

Taken together, these empirical findings indicate that low-skill students who enroll in college tend to drop out more than their higher-skilled peers, and that college dropouts tend to have trouble repaying their student loans later in life. This motivates us to study student loan policy design in a context that incorporates college dropout among students with low skill.

2.1 Student loans in the SCF

The SCF is a repeated cross-sectional survey of US households which collects information on household balance sheets every three years. We use the 2019 SCF for two reasons. First, our policy analyses will focus on the pre COVID-19 US economy. Second, the 2019 survey contains more detailed information than previous surveys for education loans.[6] In the SCF, the term "education loans" is more general than the federal student loan program. The majority of education loan balances are owed to the federal government, but not all of them are. In this paper, we focus on the federal student aid program, because loans issued through that program have terms that are a policy choice of the government. The terms of private loans are determined by market forces. Therefore, in what follows, we restrict attention to education loans for postsecondary students borrowed via federal lending programs.

To examine the composition of households with student debt in terms of college completion and repayment delinquency, we have to make several decisions of how to aggregate loan-level information within a house-

[6] "Education loans" is the term used in the SCF codebook. We use this term interchangeably with "student loans".

hold. The structure of the SCF is to record information on up to 6 education loans owed by the household. For each loan, there is a separate set of variables that record loan-specific responses to various questions about the loan. This structure serves to make the SCF very informative, but it also makes the definition of a "dropout" household or a household that is "delinquent" , variables most relevant to us, somewhat ambiguous. This is because one can construct an education loan portfolio for the household but then must decide how to map from a portfolio where none, one, or all loans may be late in repayment to a 0 or 1 indicator for the household being delinquent. Similarly, one must decide how to map from a portfolio where none, one, or all loans may be associated with a household that dropped out of the program the loan was used to pay for to a 0 or 1 indicator for the household having dropped out. Although we consider the mapping in many ways, we get similar statistics for our variables of interest.

In Table 1, we report several statistics regarding education loans in the United States, as measured in the 2019 SCF (see Table 9 in the Appendix A.2 for details on variables used). In Table 1, only loans owed by the respondent or their spouse are considered. "Not completing education" means that for all education loans owed by one of the spouses in the household, the education program was not completed. "Non-repayers" means that the household is not making payments on at least one of the education loans that they owe.

Panel A shows that, among households with positive student loan balances, 19 percent did not complete their education, 16 percent are currently enrolled or in the grace period following graduation (in which repayments are not required), and 65 percent have completed their education and are past the grace period. In terms of how the balances are distributed across these categories, 11 percent of federal education debt is held by debtors who did not complete the program the loan was used to pay for, 16 percent is held by debtors who are either currently enrolled or in the grace period after graduation, while 73 percent of the debt is held by graduates who are past the grace period. The quantity of debt owed by dropouts and those currently enrolled or in the grace period is lower than their share of the population of debtors because these two groups have not, on average, used their loans to pay for as many years of education as college graduates have. Nevertheless, dropouts are clearly a significant portion of student debtors, and they do hold a sizeable fraction of the outstanding federal student debt in the United States.

Panel B, meanwhile, shows that among households who are not enrolled or in the grace period (dropouts or graduates) the share of delinquent households who did not complete their education is much higher than the proportion of the population of debtors in this category. This reflects the difficulty of making payments on a loan without simultaneously enjoying the wage premium from a college degree.

Table 1: Education loans in the 2019 Survey of Consumer Finances

Variable (unit = percentage)	Sample	Value
Panel A. Educational attainment in the population of debtors		
Share of HH: dropouts	HH with $SL > 0$	18.35
Share of HH: enrolled or in grace period	HH with $SL > 0$	16.35
Share of HH: graduated	HH with $SL > 0$	65.30
Share of \$: dropouts	HH with $SL > 0$	10.79
Share of \$: enrolled or in grace period	HH with $SL > 0$	15.77
Share of \$: graduated	HH with $SL > 0$	73.44
Panel B. Educational attainment in the population of delinquent debtors		
Dropouts	HH who are delinquent	51.27-57.72
Graduates	HH who are delinquent	42.28-48.73

This table presents stistics from the 2019 SCF on educational loans. Panel A presents information on the composition of all households with a positive balance of educational debt, and how the aggregate balance of student debt is distributed across different types of households. Panel B presents statistics for households that are not making payments on their educational loans because they cannot afford to do so (delinquent households).

2.2 HSLS and dropouts

The HSLS is a representative panel survey of 9th-grade (high school) students in the United States beginning in 2009 (see Table 8 in Appendix A.1 for more information on the structure of the HSLS). This survey contains information on the focal student (for example their high school GPA, both in the base year at at the end of high school) and on their family (family income, parental education, etc.). Crucially, this survey also records not only high school educational outcomes but the post-secondary educational outcome of the focal student: whether they enrolled in college after high school, whether they took out student loans to pay for it, and whether they progressed in their program after enrollment. With this information, we examine the relationship between student skill (high school honors-weighted GPA) and educational outcomes, both unconditionally and conditional on other attributes that may be correlated with student skill.

In particular, using the population of high school graduates from the HSLS, we restrict our sample to students who have graduated from high school in the summer of 2013. Among this group, we additional restrict attention to observations for whom family income, the student's honors-weighted high school GPA, state of residence at the first follow-up, parental educational attainment, student race, and student ethnicity were all reported. This leaves us with a sample size of 13,700 students.[7] Using the distribution observed within this sample, we then assign students to a quintile of the high school GPA.

In Table 2, we report the persistence into the second, third, and fourth year of college is reported for each skill quintile, conditional on enrollment in the 2013-2014 academic year. Persistence rates are starkly increasing

[7]Sample counts are rounded to the nearest 10 according to NCES requirements.

in student skill.

Table 2: Persistence rates in post-secondary education by HS GPA quintiles

Quintile	Total obs.	Quintile obs.	Year 2 | Year 1	Year 3 | Year 1	Year 4 | Year 1
1	4,430	120	69	54	39
2	4,430	410	76	61	55
3	4,430	770	84	75	69
4	4,430	1,300	91	84	79
5	4,430	1,830	94	91	86

Notes: Table 2 presents persistence rates into years 2, 3, and 4, conditional on enrolling in 2013-2014 academic year at a 4-year program (year 1). Data source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study Restricted-Use Data File. Sample counts are rounded to the nearest 10 according to NCES requirements.

Having offered some unconditional moments on educational persistence, we now report several regressions to see whether these margins continue to be predictive controlling for other (potentially correlated) variables. Table 3 reports regressions where the dependent variable takes a value of 1 if the student persists to the subsequent year, conditional on being currently enrolled. Across columns, each regression conditions on being enrolled in a different academic year: the first model is for those enrolled in the first year, the second for those enrolled in the third year, and the third model reports persistence rates for those enrolled in their third year of college. For each of these three models, the controls include fixed attributes of the student (their high school GPA as a continuous variable, family income during high school, and parent educational attainment category) and attributes of the student that vary over time and are referencing the current school year for the respective model. These variables include logged hours worked, an indicator for not working, the logged balance of federal student loans, and an indicator for not having student loans.

As the findings indicate, honors-weighted high school GPA remains strongly predictive of persistence at every stage of the student's progression through post-secondary education, even when controlling for family income, parent educational attainment, hours worked, and student loan uptake. At the same time, employment patterns and student loan uptake, both along the extensive and intensive margins, are also predictive of persistence. In particular, if a student does not work this is correlated with a lower persistence probability, but if they do work then persistence probability is decreasing in hours, so that working to much is worse for persistence than not working at all. For federal student loan balances, not having any student loans increases one's probability of persistence into years 2 and 3, but not into year 4, even controlling for student skill. At the same time, the balance of student loans positively predicts persistence, so that a student with a high balance of loans is as likely to persist as one without any loans at all.

Table 3: Predicting continued enrollment in subsequent year, conditional current enrollment

Variable	**Year 2 | Year 1**	**Year 3 | Year 2**	**Year 4 | Year 3**
Student HS GPA (honors-weighted)	0.116***	0.0886***	0.0558***
	(0.0152)	(0.0134)	(0.0144)
Average family income during HS	2.04e-08	1.26e-08	1.88e-08
	(1.81e-08)	(1.76e-08)	(1.68e-08)
At least 1 parent AA+	0.00588	0.0315	0.0414
	(0.0293)	(0.0336)	(0.0286)
Log hours worked current year	-0.0294**	-0.0380***	-0.0273**
	(0.0145)	(0.0133)	(0.0116)
Not working current year	-0.0997***	-0.104***	-0.0995***
	(0.0384)	(0.0343)	(0.0340)
Log cumulative federal student loan current year	0.0762***	0.0400**	0.0125
	(0.0223)	(0.0167)	(0.0138)
No federal student loan current year	0.639***	0.369**	0.134
	(0.193)	(0.156)	(0.131)
Constant	-0.0423	0.348**	0.670***
	(0.210)	(0.175)	(0.158)
r^2	0.08	0.09	0.07
Observations (rounded)	4430	3980	3670

Notes: The dependent variable in each column is an indicator for persisting to the subsequent year of college, conditional on being enrolled in the current year. The independent variables are a continuos measure of student skill (honors-weighted HS GPA), the average family income observed for the focal student in the HSLS during their time in high school, the education category of their family (high education means having at least one resident parent with an associate's degree), the log of the hours they usually worked per week in the current year (this variable is set to 0 if they do not work), an indicator for not working while enrolled during the current year, the log of their student loan balance in the current year (set to 0 if they have no student loans), and an indicator for having a cumulative balance of 0 for federal student loans. Standard errors in parentheses: $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$. Data source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study Restricted-Use Data File. Sample counts are rounded to the nearest 10 according to NCES requirements.

3 Model

Our model economy builds on Krueger and Ludwig (2016), Chatterjee and Ionescu (2012), and Luo and Mongey (2019). We enrich Krueger and Ludwig (2016)'s general equilibrium life cycle model with college choice by incorporating over-optimism about graduation likelihood. We base our modeling assumptions about over-optimism on the empirical findings of Stinebrickner and Stinebrickner (2012). We also incorporate endogenous and exogenous college dropouts and key features of the U.S. federal student loan program into our model economy. These features are based on the models in Chatterjee and Ionescu (2012) and Luo and Mongey (2019), respectively.

3.1 Overview of the model environment

Time is discrete and runs forever. A period is one year. For ease of exposition, although we will compute transition paths, we omit time subscripts. Let j denote the age of agents. Agents start making decisions when they turn 18 ($j = 1$ in the model). At the beginning of each period, 18-year-olds make a one-time decision of whether to go to college or not. This decision will be based on their idiosyncratic permanent productivity, idiosyncratic AR(1) productivity, initial net assets, and high school GPA (HS GPA). The permanent productivity component and AR(1) productivity component are denoted by ν and η, respectively. The initial net assets are denoted by a and are determined by a one-time inter vivos transfer from the agents' parents. HS GPA is denoted by e and is a one-time draw that depends on the parents' education. HS GPA determines the agents' annual probabilities of continuing their college education (i.e., annual probabilities of not dropping out, given by $p_g(j, e)$, see Section 3.2), the annual effort cost for a college education $\lambda(e)$, and subsidies for college from the government and private sources.

When making the college entrance decision, agents are over-optimistic about their likelihood of graduation. Let $\hat{p}$ denote the minimum probability with which agents believe they will continue their education in each year of college. As mentioned above, the true annual probability of continuing a college education is given by the function $p_g(j, e)$. The agents are overly-optimistic because they use the maximum of $p_g(j, e)$ and $\hat{p}$ to compute the expected value from going to college. This does not imply that the extent of over-optimism is the same across students. For $p_g(j, e) < \hat{p}$, a lower $p_g(j, e)$ would imply higher over-optimism. Agents correctly understand their environment if $\hat{p} = 0$. We assume that $\hat{p}$ is the same across agents. This assumption is motivated by Stinebrickner and Stinebrickner (2012), who find little variation in expected academic performance across HS GPA groups.

Once they enter college, agents incorporate their true probability of graduation. Agents may choose to dropout of college after the first year of college (endogenous dropout) or may dropout of college with an annual probability of $p_d(j, e) = 1 - p_g(j, e)$ (exogenous dropout). The exogenous dropout captures dropouts due to differences in academic ability.[8]

In our model economy, college lasts 4 years. The annual cost of college (tuition and fees) is denoted by κ. As in Chatterjee and Ionescu (2012) and Krueger and Ludwig (2016), we do not include room and board as a mandatory expense in our model. Section 5.2 discusses the rationale for this modeling choice in more detail. Agents in college work part time ℓ_{pt}, and hence, forego earnings from full time work. Additionally, they incur the effort cost $\lambda(e)$. The benefit of graduating from college is higher labor earnings (see Section

[8]An alternative modeling assumption would be to assume that it takes longer for agents to realize their true probabilities of continuing in college, i.e., agents are over-optimistic for a longer time period. Our assumption is more conservative about the extent of over-optimism and the paternalistic role for government.

3.5). If agents dropout of college at any time, then there is no benefit from attending college.

To pay for college, students may use inter vivos transfers from their parents, use subsidies from the government and private sources, use earnings from part time work, or borrow from the federal student loan program. The federal student loan program is characterized by a cumulative student loan limit $\bar{A}$ and a constant student loan interest rate r_{SL}, which is assessed starting from the year after the age of college graduation ($j > 4$). That is, there is an interest free grace period for student loans for the duration of college. $\bar{A}$ is our measure of access to federal student loans and our policy variable of interest. Student loan payments are required to begin after college graduation, and the loan balance is expected to be paid off in at most T_{SL} years.

After the age of college graduation, an agent with a positive student loan balance may be either a college graduate or a dropout. At this point in the life cycle, student debtors must make a decision about whether to make their required loan payment, in addition to a consumption-savings decision. If agents choose to not make a required payment on their student loan, they are considered delinquent, and their disposable income above $\bar{y}$ is garnished at the rate τ_g. Furthermore, the agents incur a penalty cost equal to ϕ fraction of the missed payment for the particular year. This penalty captures collection fees. The penalty fee and the missed payment are added to the outstanding balance for the next period. This modeling structure is similar to the current U.S. system (see Yannelis (2020) for more institutional details on student loan delinquency and penalties). It is important to note that, in our model, it is not possible to default on a student loan and have the outstanding debt being written off. Instead, missing a payment leads to punitive confiscation of income. This is consistent with the U.S. federal student loan system, where student loans are eventually classified as defaulted loans but are almost never discharged.[9]

All agents have a child at age j_f. The child will grow up and leave the household j_a years after birth. At the beginning of the period when the child leaves the household, the parent may make an inter vivos transfer of wealth to their child, after observing the child's HS GPA. Furthermore, when making the transfer, the parent is also over-optimistic about the likelihood of their child's graduation likelihood. Agents retire at age j_r. Upon retirement, agents stop working and receive social security. All agents survive each period with conditional survival probability ψ_j and live for a maximum of J periods.

The government, in addition to running the federal student loan program, providing public subsidies for college education, and providing social security, incurs a government consumption cost (a fixed fraction of GDP). These expenditures are financed through consumption taxes with a flat rate and income taxes that are progressive. A final goods firm operates a Cobb-Douglas production technology, where the inputs are capital and efficiency units of labor.

[9] In reality, for federal student loans in the United States, 270 days of missed payments is necessary before the loan is classified as being defaulted on and the associated penalties are applied. Our model period is 1 year, so we cannot distinguish in the model between a delinquency and a default. In the public-use SCF, the date variables are suppressed, so we also cannot compute days or months since the last payment made on loans. This makes it difficult to identify in the data which households are missing payments (delinquent) and which ones are missing many payments (defaulting). We use the term delinquent throughout this paper to refer to the less stringent category, as in Luo and Mongey (2019).

3.2 College dropouts

Dropping out of college can be exogenous or endogenous. As mentioned above, $p_d(j,e)$ denotes the true annual exogenous probability of dropping out of college, which is decreasing in the academic ability (skill) of the agent. This modeling choice is also based on empirical findings in Stinebrickner and Stinebrickner (2012); in that paper, the authors argue that it is heterogeneity in ability, rather than heterogeneity in effort, that drives the college dropout decision. For example, even for students in the same major who put in the same hours of study, they find significant differences in academic performance. We capture this feature of the data through the exogenous dropout probability function given by

$$p_d(j,e) = (1-p(e))\rho_d(e)^{j-1}, \tag{1}$$

where $p(e)$ determines the probability of continuing in college and $\rho_d(e)$ determines the persistence of the probability of dropping out of college conditional on college year. To illustrate, if $p(e)$ is low, the student is more likely to drop out of college. If $\rho_d(e)$ is low, the college dropout probability is less persistent with each year of college, and hence, the student is less likely to drop out of college the longer they continue their education. This functional form is a parsimonious way to capture the fact that most college dropouts leave in the first year of college and students are less likely to drop out of college with each additional year of education. The probability of continuing in each year of college $p_g(j,e)$, is equal to $1-p_d(j,e)$.

3.3 Student loan payments

As mentioned above, agents are expected to make payments starting at age $j=5$ (age after college graduation). The loans are expected to be paid off in at most T_{SL} years. Equation 2 specifies the full repayment function $\rho_R(a,j)$. If there is an outstanding student loan ($a<0$) and if j is still within the standard repayment period ($4<j\leq T_{SL}+4$), the loan is amortized with an interest rate of r_{SL} (first case in equation 2). If there is an outstanding loan and if the standard repayment period has expired ($j>T_{SL}+4$), then the outstanding principal plus interest is due. If $a>0$, then there is no student loan, and hence, no minimum payment.

$$\rho_R(j,a) = \begin{cases} -\dfrac{r_{SL}}{1-(1+r_{SL})^{-(TSL+5-j)}}a & \text{if} \quad a<0 \quad \text{and} \quad 4<j\leq T_{SL}+4 \\ -(1+r_{SL})a & \text{if} \quad a<0 \quad \text{and} \quad j>T_{SL}+4. \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Instead of repayment, suppose agents choose to default. As mentioned above, loans in default are not discharged. Instead, the agents' disposable income income above $\bar{y}$ are garnished at the rate τ_g. This leads to a partial repayment function in default given by

$$\rho_D(a,y) = \min[\tau_g \max[y-T(y)-\bar{y},0], -(1+r_{SL})a], \tag{3}$$

where $-(1 + r_{SL})a$ is the maximum amount that can be garnished.

3.4 Preferences

An agent's utility depends on total household consumption c, the agent's age j (which also determines whether or not they are raising a child), and their education status $s \in \{h, \ell\}$. h refers to an high-education agent who is either enrolled in college or is a college graduate, and ℓ refers to a low-education agent who did not go to college or who dropped out of college. Together with j, s indicates if an agent is in college or not. Agents have CRRA utility over per-capita household consumption with a relative risk aversion given by σ. When the child lives with the agent, $j \in \{j_f, .., j_f + j_a - 1\}$, the child will be included in total household consumption with an adult equivalence parameter ζ. Agents who go to college ($s = h$ and $j \in \{1, 2, 3, 4\}$), are subject to the effort cost $\lambda(e)$. The utility function is given by

$$U(c, j, s) = \frac{\left(\frac{c}{1+\zeta\, 1_{j \in \{j_f,..,j_f+j_a-1\}}}\right)^{1-\sigma}}{1-\sigma} - \lambda(e) 1_{s=h \text{ and } j \in \{1,2,3,4\}}. \tag{4}$$

3.5 Income

Income depends on age, education, earnings productivities, and net assets, summarized by the tuple (j, s, ν, η, a). $a > 0$ indicates positive net assets, which bear a savings interest rate r, and $a < 0$ indicates outstanding student loans, which bear a student loan interest rate r_{SL}. The income function is given by:

$$y_{j,s,\nu,\eta,a} = \begin{cases} s = h & \begin{cases} w\epsilon_{j,\ell}\nu\eta\ell_{pt} & \text{if } j = 1 \\ w\epsilon_{j,\ell}\nu\eta\ell_{pt} + r\left[\max(a, 0) + Tr_j\right] & \text{if } 1 < j \leq 4 \\ w\epsilon_{j,h}z\nu\eta + r\left[\max(a, 0) + Tr_j\right] & \text{if } 4 < j < j_r \\ ss_{h,\nu} + r\left[\max(a, 0) + Tr_j\right] & \text{if } j \geq j_r \end{cases} \\ s = l & \begin{cases} w\epsilon_{j,\ell}\nu\eta & \text{if } j = 1 \\ w\epsilon_{j,\ell}\nu\eta + r\left[\max(a, 0) + Tr_j\right] & \text{if } 1 < j < j_r \\ ss_{\ell,\nu} + r\left[\max(a, 0) + Tr_j\right] & \text{if } j \geq j_r, \end{cases} \end{cases} \tag{5}$$

where $\epsilon_{j,s}$ is a deterministic life cycle productivity component that depends on education and z is a college premium adjustment parameter that ensures the model matches the observed college premium in the data. w is the wage rate and Tr_j is accidental bequests, which depends on age. Accidental bequests are a consequence of the assets of the deceased.

When agents first enter the labor market at 18 ($j = 1$), their only source of income is labor earnings. If an agent chooses to go to college ($s = h$), they work part time and receive $w\epsilon_{j,\ell}\nu\eta\ell_{pt}$ in labor earnings. Because this agent is still in college, their life cycle productivity component and their AR(1) productivity component are drawn from the distribution for an agent without a college degree. If a working-age agent does not go to college or drops out of college ($s = \ell$), they work full time and receive $w\epsilon_{j,\ell}\nu\eta$ in labor earnings. If a working-age agent is a college graduate, they also work full time and their labor earnings is given by $w\epsilon_{j,h}z\nu\eta$. In this case, the deterministic life cycle component and the AR(1) productivity component are drawn from the distribution for an agent with a college degree. When agents retire at age j_r, they do not work but receive social security $ss_{s,\nu}$. The level of social security payments an agent receives depends on both their education and the permanent productivity component they draw (refer to equation 25).

When agents are 18, they receive an inter vivos transfer from their parents. As the interest income is collected by their parents, these agents do not receive interest income on their inter vivos transfer. However, after the first year ($j > 1$), the agents' income, in addition to labor earnings or social security, also includes any interest from positive net assets and accidental bequests ($r[\max[a, 0] + Tr_j]$).

3.6 Consumer problems before college graduation age (j ≤ 4)

Let the tuple (ν, η, a, e) summarize the agent's idiosyncratic permanent productivity, AR (1) productivity, net assets, and HS GPA. The 18-year-old ($j = 1$) solves the college entrance problem given by

$$\hat{W}(\nu, \eta, a, e) = \max_{\hat{d}_s \in \{0,1\}} (1 - \hat{d}_s)V(1, \ell, \nu, \eta, a) + \hat{d}_s \hat{V}(1, h, \nu, \eta, a, e), \tag{6}$$

where $\hat{d}_s \in \{0, 1\}$ is the discrete college entrance decision, $V(1, \ell, \nu, \eta, b)$ is the value of not going to college, and $\hat{V}(1, h, \nu, \eta, a, e)$ is the over-optimistic value of going to college.

The value of not going to college (as well as the value of dropping out) for $j \leq 4$ is given by

$$V(j, \ell, \nu, \eta, a) = \max_{c,a'} U(c, j, \ell) + \beta\psi_j E_{\eta'|\ell,\eta} V(j + 1, \ell, \nu, \eta', a') \tag{7}$$

$$\begin{aligned}
&s.t.\\
&(1 + \tau_c)c + a' = y_{j,\ell,\nu,\eta,a} + a + Tr_j - T(y_{j,\ell,\nu,\eta,a})\\
&a' \geq 1_{a<0}a\\
&c \geq 0,
\end{aligned}$$

where β is the discount factor, τ_c is the consumption tax rate, and $T(y_{j,\ell,\nu,\eta,a})$ is the income tax function (see Section 3.8). As in Krueger and Ludwig (2016), the AR(1) productivity process is different for those with and without a college degree. Therefore, this agent draws their next period AR(1) productivity from the expectation operator that depends on ℓ, in addition to η. Agents who do not go to college will not take on any student loans. Furthermore, for tractability, we abstract from other forms of borrowing. Hence, in equilibrium, for agents that do not go to college, net assets are always weakly positive ($a \geq 0$).

Equation 7 is also the value function for agents that drop out of college. These agents may have outstanding student loans ($a < 0$). In that case, by construction, the agent is not assessed interest until $j > 4$, and is allowed to roll over their debt. However, these agents cannot take on more debt because they are no longer enrolled in college. Thus their borrowing limit, if they have outstanding loans, is equal to $-a$. These agents can pay down their debt as much as they want to and as they pay down their debt, their borrowing limit will shrink.

As mentioned above, $\hat{V}(j,h,\nu,\eta,a,e)$ is the over-optimistic value of college. Hence, when computing the value of college, agents use the maximum of the true probability of continuing in each year of college $p_g(j,e)$ and the over-optimistic probability of continuing in each year of college $\hat{p}$. The over-optimistic value of college for $j = 1,2,3$ is given by

$$
\begin{aligned}
&\hat{V}(j,h,\nu,\eta,a,e) = \max_{\hat{c},\hat{a}'} U(c,j,h) + \qquad (8)\\
&\beta\psi_j E_{\eta'|\ell,\eta}\Big[\max[p_g(j,e),\hat{p}]\max[\hat{V}(j+1,h,\nu,\eta',\hat{a}',e), V(j+1,\ell,\nu,\eta',\hat{a}')] + \\
&(1-\max[p_g(j,e),\hat{p}])V(j+1,\ell,\nu,\eta',\hat{a}')\Big]\\
&s.t.\\
&(1+\tau_c)\hat{c} + \hat{a}' + (1-\theta(e)-\theta^{pr}(e))\kappa = y_{j,h,\nu,\eta,a} + a + Tr_j - T(y_{j,h,\nu,\eta,a})\\
&\hat{a}' \geq -\frac{\bar{A}[(1-\theta(e)-\theta^{pr}(e))\kappa + \bar{c}]j}{4}\\
&\hat{c} \geq 0,
\end{aligned}
$$

where $\theta(e)$ and $\theta^{pr}(e)$ are the share of tuition and fees that are paid for by public and private subsidies, and are a function of HS GPA. This source of aid is distinct from student loans and is not repayed, like grants and scholarships. $\bar{c}$ is the amount that can be borrowed for room and board expenses while in college. These agents mistakenly believe that they may exogenously drop out of college with probability $1 - \max[p_g(j,e),\hat{p}]$, where the max operator reflects the potential over-optimism of agents. These agents can also choose to drop out after the first year of college captured by the expression $(\max[\hat{V}(j+1,h,\nu,\eta',a',e), V(j+1,\ell,\nu,\eta',a')])$. Furthermore, these agents can borrow up to a student loan limit equal to $\frac{\bar{A}[(1-\theta(e)-\theta^{pr}(e))\kappa + \bar{c}]j}{4}$. $\bar{A}$ is the number of years of net tuition and fees plus room and board expenses that the student loan limit can cover. For example, if $\bar{A}$ is equal to four, then the student loan limit is equal to four years of net tuition and fees plus room and board. The multiplier $\frac{j}{4}$ is an adjustment for the fact that the total student loan limit increases with each year of college. This imposes hat students can only borrow to pay for net tuition and fees plus room and board.

When agents reach the final year of college ($j = 4$), their over-optimistic continuation value, if they are graduating college, will be based on $E_{\eta'|h,\eta}$ rather than $E_{\eta'|\ell,\eta}$. Furthermore, there will be no endogenous dropout decision in the continuation value at this age, because in the next period the agent will have graduated from college. The rest of the value function for the final year of college remains unchanged (refer to Appendix B, equation 20).

It is important to emphasize that when agents make the college entrance decision (equation 6), agents are over-optimistic and will use the over-optimistic value of college from equations 8 and 20 to compute their expected value. However, as mentioned above, we assume that agents learn their true dropout probabilities in the first year of college. Therefore, once students are in college, i.e., in the simulation of the model, the consumption-savings and dropout decisions are based on the following value function in college for $j = 1, 2, 3$

$$
\begin{aligned}
&V(j,h,\nu,\eta,a,e) = \max_{c,a'} U(c,j,h) + \qquad (9)\\
&\beta\psi_j E_{\eta'|\ell,\eta}\Big[p_g(j,e)\max[V(j+1,h,\nu,\eta',a',e), V(j+1,\ell,\nu,\eta',a')] +\\
&(1-p_g(j,e))V(j+1,\ell,\nu,\eta',a')\Big]\\
&s.t.\\
&(1+\tau_c)c + a' + (1-\theta(e)-\theta^{pr}(e))\kappa = y_{j,h,\nu,\eta,a} + a + Tr_j - T(y_{j,h,\nu,\eta,a})\\
&a' \geq -\frac{\bar{A}[(1-\theta(e)-\theta^{pr}(e))\kappa + \bar{c}]j}{4}\\
&c \geq 0.
\end{aligned}
$$

The only difference between this value function and the over-optimistic value function (represented by (8)) is that agents use the true probabilities of continuing in college in computing their expected value from college. Again, when these agents reach the final year of college ($j = 4$), if they are graduating from college, their continuation of value will be based on $E_{\eta'|h,\eta}$ rather than $E_{\eta'|\ell,\eta}$. The rest of their value function for the final year of college remains unchanged (refer to Appendix B, equation 21).

3.7 Consumer problems after college graduation age (j $\geq$ 5)

Agents start to repay any student loans the year after college graduation age, regardless of whether they complete college or drop out. In the United States, federal student loans typically have a six month grace period after graduation, in which repayment does not need to be made. Since our model period is one year, the fact that repayment starts one period after graduation is consistent with the current policy. For $j \geq 5$, agents choose between repayment and default in every period. The idiosyncratic state of an agent for $j \geq 5$ and $j \neq j_f + j_a$ is summarized by the tuple (j, s, ν, η, a). The agent's value function is given by

$$
\begin{aligned}
&V(j,s,\nu,\eta,a) = \max_{d_f\in\{0,1\}} (1-d_f)V^R(j,s,\nu,\eta,a) + d_f V^D(j,s,\nu,\eta,a) \qquad (10)\\
&s.t.\\
&d_f = 0 \qquad \text{if} \qquad a \geq 0,
\end{aligned}
$$

where $d_f \in \{0, 1\}$ is the student loan repayment decision, $V^R(j, s, \nu, \eta, a)$ is the value of repayment, and $V^D(j, s, \nu, \eta, a)$ is the value of default. The agent may default if they have an outstanding student loan. The value of repayment for $j \geq 5$ and $j \neq j_f + j_a$ is given by

$$V^R(j,s,\nu,\eta,a) = \max_{c,a'} U(c,j,s) + \beta\psi_j E_{\eta'|s,\eta} V(j+1,s,\nu,\eta',a') \tag{11}$$
$$s.t.$$
$$(1+\tau_c)c + a' = y_{j,s,\nu,\eta,a} + a + 1_{\{a<0\}} r_{SL} a + Tr_j - T(y_{j,s,\nu,\eta,a})$$
$$a' \geq 1_{a<0}(1+r_{SL})a + \rho_R(j,a)$$
$$c \geq 0.$$

The agent must make a minimum payment of $\rho_R(j,a)$ (see Section 3.3). Thus their borrowing limit, if they have an outstanding balance, decreases to $-a(1+r_{SL}) - \rho_R(j,a)$. The value of default for $j \geq 5$ and $j \neq j_f + j_a$ is given by

$$V^D(j,s,\nu,\eta,a) = U(c,j,s) + \beta\psi_j E_{\eta'|s,\eta} V(j+1,s,\nu,\eta',a') \tag{12}$$
$$s.t.$$
$$(1+\tau_c)c = y_{j,s,\nu,\eta,a} + Tr_j - T(y_{j,s,\nu,\eta,a}) - \rho_D(a, y_{j,s,\nu,\eta,a})$$
$$a' = (1+r_{SL})a + \rho_D(a, y_{j,s,\nu,\eta,a}) - \phi \max[\rho_R(j,a) - \rho_D(a, y_{j,s,\nu,\eta,a}), 0].$$

In the case of non-repayment, agents do not make a consumption-savings decision. Instead, these agents have their income garnished leading to a partial payment of $\rho_D(a, y_{j,s,\nu,\eta,a})$. Therefore, they consume whatever remains from their disposable income plus accidental bequests after making the partial payment. As mentioned above, ϕ 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 (minus any partial payment).

Agents at age $j = j_f + j_a$ may make a one-time inter-vivo transfer to their child in addition to the default and consumption-savings decisions. At this age, the child becomes independent and leaves the parent's household. Agents make their decisions after observing their child's HS GPA. The agent's value function is given by

$$V(j,s,\nu,\eta,a) = \sum_e \pi(e|s)\big[\max_{d_f \in \{0,1\}} (1-d_f)V^R(j,s,\nu,\eta,a,e) + d_f V^D(j,s,\nu,\eta,a,e)\big] \tag{13}$$
$$s.t.$$
$$d_f = 0 \qquad \text{if} \qquad a \geq 0,$$

where $\pi(e|s)$ is the child's HS GPA probability function that depends on the parent's education. The value of repayment for $j = j_f + j_a$ is given by

$$
\begin{aligned}
&V^R(j,s,\nu,\eta,a,e) = \max_{c,a',b} U(c,j,s) + \beta\psi_j E_{\eta'|s,\eta} V(j+1,s,\nu,\eta',a') && (14)\\
&+ \beta_c E_{\nu'|\nu} E_{\eta'|\ell} \hat{W}(\nu',\eta',b,e)\\
&s.t.\\
&(1+\tau_c)c + a' + b = y_{j,s,\nu,\eta,a} + a + 1_{\{a<0\}} r_{SL} a + Tr_j - T(y_{j,s,\nu,\eta,a})\\
&a' \geq 1_{a<0}(1+r_{SL})a + \rho_R(j,a)\\
&c \geq 0\\
&b \geq 0,
\end{aligned}
$$

where b is the inter vivos transfer, $\hat{W}(\nu',\eta',b,e)$ is the child's value function, and β_c disciplines the intensity of parental altruism towards the child. The child's permanent productivity ν' depends on the parent's permanent productivity, which leads to potential persistence in intergenerational earnings. Note that this can be interpreted to represent exogenous transfers of human capital across generations of the same family, a simplification of a more complex skill investment decision.[10] The child's AR(1) productivity η' is drawn from the stationary distribution for an agent without a college degree. If the agent chooses to default instead of repay, then they will not make an inter vivos transfer to their child ($b = 0$) and are subject to the standard default rules as in equation 12 (refer to Appendix B, equation 22).

3.8 Production and government budget constraint

The production function is Cobb-Douglas given by

$$K^\alpha (ZL)^{1-\alpha}, \tag{15}$$

where K is aggregate capital stock, Z is aggregate labor productivity, L is total efficiency units of labor, and α is the capital share. This production function assumes that an efficiency unit of labor from a college graduate is perfectly substitutable with an efficiency unit of labor from a worker without a college degree. Because the general equilibrium effects of our policy analyses are small, this assumption will not affect our results. The capital stock depreciates at rate δ. The representative firm rents capital at an interest rate $r + \delta$ and hires workers at the wage rate w. First order conditions from the firm's optimization problem leads to the following standard conditions

$$
\begin{aligned}
r =& \alpha K^{\alpha-1}(ZL)^{1-\alpha} - \delta && (16)\\
w =& (1-\alpha)K^\alpha L^{-\alpha} Z^{1-\alpha}. && (17)
\end{aligned}
$$

The government collects consumption and income taxes. The income tax function following Heathcote et al.

[10] In fact, in this model, there are two ways that human capital in this model reflects family background: the child's high school GPA, e, and the permanent component of the child's earnings shock, ν'. Both of these depend on parent characteristics.

(2017) is given by

$$T(y) = y - \gamma y^{1-\tau_p}, \tag{18}$$

where τ_p governs the tax progressivity of income taxes and γ is used to balance the government budget in every period. The government uses its tax revenue to finance government consumption, social security, public tuition subsidies, and the federal student loan program.

Lastly, refer to Appendix B.2 for the equilibrium definition.

4 Calibration

Given the computationally demanding nature of our model, we calibrate as many parameters as possible outside the model equilibrium and calibrate the remaining parameters jointly to target moments in the US economy. Tables 4 and 5 present parameters determined outside of the model equilibrium and Table 6 presents parameters calibrated to target moments in the US economy.

Table 4 presents parameters calibrated outside of the model equilibrium that are related to the federal student loan program and college education. We set the aggregate federal student loan limit to 1.737, which is the aggregate limit for a dependent student since 2012, normalized by the average annual net tuition and fees plus room and board. Hence, the estimate of 1.737 implies that the limit is sufficient to pay for 1.737 years of average annual net tuition and fees plus room and board. Here, we use the borrowing limit for dependent students rather than independent students. The borrowing limit for dependents is significantly lower that it is for independents, and we use the former for two reasons. First, most undergraduate students in the United States are dependents (NCES (2020)). Second, as we will show in Section 5.2, the baseline model with student loan limits equal to those of dependent students rationalizes student loan balances among fourth year students. The model with a student loan limit equal to that of an independent student significantly overestimates outstanding balances among fourth year students.

We set the student loan interest rate r_{SL} to 0.058, which is the average real interest rate on federal student loans from 1992-2019 (SCF). We set the maximum number of years to repay the student loan T_{SL} to 20. In the US, those with student loans may choose between a standard repayment plan of ten years and an income based repayment plan, which varies from 10-25 years (CRS (2019)). We do not view the abstraction from these repayment plans in our model as undermining our results. This is because agents in our model can choose to repay their loan at a faster rate, for e.g., in ten years like the standard repayment plan, or at a slower rate, for e.g., in 20 years like the income based repayment plan. As for the garnishment rate conditional on default, τ_g, we set it to 15 percent, which is the current rate of garnishment for federal student loans.

To calibrate shares of students by HS GPA quintile given parental education, i.e., $\pi(e|s)$, we use the HSLS. As Table 4 reports, parents with a college education are more likely to have children with higher high school GPAs. To calibrate shares of college tuition subsidized by the government $\theta(e)$ and shares of college tuition subsidized private beneficiaries $\theta^{pr}(e)$, we combine estimates from HSLS with estimates from Krueger and Ludwig (2016). First, we compute shares of tuition subsidized by the government or private sources

Table 4: Parameters calibrated outside of model equilibrium, part 1

Parameter		Description	Value
Panel A. Student loan program			
$\bar{A}$	Limit	CRS (2019)	1.737
r_{SL}	Interest rate	SCF average 1992-2019	0.058
T_{SL}	Maximum years to repay	CRS (2019)	20
τ_g	Garnishment rate	Yannelis (2020)	0.150
Panel B. Education			
$\pi(e_1 \vert s = h)$			0.140
$\pi(e_2 \vert s = h)$			0.170
$\pi(e_3 \vert s = h)$			0.190
$\pi(e_4 \vert s = h)$			0.230
$\pi(e_5 \vert s = h)$	Student share by HS-GPA | Parental education	HSLS	0.260
$\pi(e_1 \vert s = \ell)$			0.240
$\pi(e_2 \vert s = \ell)$			0.220
$\pi(e_3 \vert s = \ell)$			0.190
$\pi(e_4 \vert s = \ell)$			0.190
$\pi(e_5 \vert s = \ell)$			0.170
$\theta(e_1)$			0.470
$\theta(e_2)$			0.338
$\theta(e_3)$	Public tuition subsidy | HS GPA	HSLS + Krueger and Ludwig (2016)	0.358
$\theta(e_4)$			0.379
$\theta(e_5)$			0.501
$\theta^{pr}(e_1)$			0.470
$\theta^{pr}(e_2)$			0.338
$\theta^{pr}(e_3)$	Private tuition subsidy | HS GPA	HSLS + Krueger and Ludwig (2016)	0.358
$\theta^{pr}(e_4)$			0.379
$\theta^{pr}(e_5)$			0.501

Panel A reports the policy parameters of the federal student loan program. Panel B reports the conditional probability of drawing high school GPA e given parental education s, the public tuition subsidy, and the private tuition subsidy (i.e., grants and scholarships). Data sources are provided in the second column.

in the HSLS by HS GPA quintile. In the HSLS, we cannot distinguish if the subsidy was received from the government or a private source, so we proceed as follows.Krueger and Ludwig (2016) estimate that government subsidies pay for 38.8 percent of total tuition and private subsidies pay for 16.6 percent of total tuition. This implies the government's share of tuition subsidies is 70 percent and private beneficiaries' share of tuition subsidies is 30 percent. To assign values to $\theta(e)$, we multiply the total share of tuition subsidized by the government or private entities by 0.7, and to assign $\theta^{pr}(e)$ we multiply it by 0.3.

Table 5 presents the more standard parameters calibrated outside of the model equilibrium. The child bearing age j_f is set to 13, which implies agents have a child when they turn 30. The number of years before the child moves out j_a is set to 18, which implies agents move out and make the college entrance decision when they turn 18. The retirement age j_r is set such that agents retire at 65 and the maximum life span J is set such that agents live for at most 100 years. The adult equivalence scale ζ is set to 0.3, following Krueger and Ludwig (2016). As in Krueger and Ludwig (2016), we set conditional survival probabilities ψ_j for $j = 1, ..., j_f + j_a - 1$ to one. This avoids there being children without parents in the model. We do not view this assumption as a major concern because young working-age consumers have high survival probabilities. For $j \geq j_f + j_a$ (i.e., for agents starting at age 48), we use estimates from the Social Security Administration life tables. For deterministic life cycle productivities conditional on education $\epsilon_{j,s}$, we use estimates from Conesa et al. (2018).

We set the relative risk aversion parameter σ to 2, which is standard in the macro literature. We set the capital share parameter α to 0.36, which implies a labor share of 64 percent. The depreciate rate δ is set to 0.076, following Krueger and Ludwig (2016). The part time work hours for students enrolled in college is set to 0.338, which is the average weekly hours from the HSLS expressed as a fraction of a 100-hour weekly time endowment.

We discretize the AR(1) earnings processes using the Tauchen (1986) method with 13 grid points. For the persistence and variance parameters that depend on college education, we use estimates from Krueger and Ludwig (2016). We set the intergenerational persistence of the permanent productivity process ρ_ν to 0.4, which is consistent with the range of estimates for intergenerational persistence of income (see Chetty, Hendren, Kline, and Saez (2014) and discussion therein).

The consumption tax rate τ_c is set to 5 percent (Krueger and Ludwig (2016)), the progressivity of the income tax function τ_p to 0.181 (Heathcote et al. (2017)), and government consumption, g, as a share of GDP, to 0.17 (BEA).

Table 6 presents the remaining 26 parameters calibrated jointly in equilibrium to target 26 moments of the US economy. It is important to emphasize that although we discuss the most significant one to one relationship between each parameter and target moment, the parameters are calibrated jointly and each parameter can affect all targets moments. A key parameter in our model is the irrational minimum graduation probability $\hat{p}$, which determines the extent of over-optimism. We calibrate it such that the expected graduation rate is 86 percent (Stinebrickner and Stinebrickner, 2012). Note that, in the model, this expected graduation rate exceeds the realized graduation rate of 71.4 percent by a significant amount.

Other important parameters related to college are the parameters that determine the true likelihood of graduation ($p(e)$ and $\rho_d(e)$) and college effort cost $\lambda(e)$. In our model, $p(e)$ captures academic ability to graduate college and depends on the HS GPA quintile. To calibrate $p(e)$, we target probabilities of enrolled students continuing to the second year of college conditional on their HS GPA quintile. Although we can directly estimate the unconditional probabilities from the HSLS, we do not use these probabilities as direct targets. Instead we control these probabilities for family income and parental education to isolate the impact of academic ability captured by HS GPA on graduation likelihood. In Appendix C.1, Table 10 reports how the persistence to second year of college conditional on enrollment is predicted by student skill quintile (with the highest quintile dropped), family income, and parent education. In the model, we run the same regression as in Table 10, and calibrate $p(e)$ to match coefficients for the first four HS GPA quintiles and the constant

Table 5: Parameters calibrated outside of model equilibrium, part 2

Parameter		Description	Value
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
ζ	Adult equivalence	Krueger and Ludwig (2016)	0.300
ψ_j	Survival probability	Krueger and Ludwig (2016) + SSA	
$\epsilon_{j,s}$	Deterministic life cycle productivity	Conesa et al. (2018)	
Preferences & technology			
σ	Risk aversion	Standard in macro literature	2
α	Capital share	Labor share = 0.64	0.360
δ	Depreciation rate	Krueger and Ludwig (2016)	0.076
ℓ_{pt}	Part time work hours while in college	HSLS average	0.338
Earnings processes			
$\rho_{\eta\|h}$	Persistence AR(1) | college graduate	Krueger and Ludwig (2016)	0.969
$\rho_{\eta\|n}$	Persistence AR(1) | in college or not college graduate		0.928
$\sigma^2_{\eta\|h}$	Variance AR(1) | college graduate		0.010
$\sigma^2_{\eta\|n}$	Variance AR(1) | in college or not college graduate		0.019
ρ_ν	Persistence permanent earnings	Intergenerational persistence = 0.4	0.400
Government			
τ_c	Consumption tax rate	Krueger and Ludwig (2016)	0.050
τ_p	Income tax progressivity	Heathcote et al. (2017)	0.181
g	Government consumption	Share of GDP (BEA)	0.170

term.

To calibrate $\rho_d(e)$, which determines the persistence of college dropout probabilities, we use the share of students that survive to the fourth year of college (senior year) conditional on surviving to the second year of college (sophomore) by their HS GPA quintile. For example, 0.570 in Table 6 implies that 57 percent of students continue their education to the fourth year of college conditional on still being enrolled in college in their sophomore year and being in the first quintile of HS GPA. We calibrate the effort cost $\lambda(e)$ to target college enrollment rates by HS GPA quintile. College enrollment rate increases from 9 percent for students in the first HS GPA quintile to 80 percent for students in the fifth HS GPA quintile.

We calibrate aggregate labor productivity Z such that GDP per capita for the population 18 and over is 1 in the model. The discount factor β is calibrated to target a capital to output ratio of 3. For annual tuition and fees κ and annual room and board $\bar{c}$, we use the annual averages from **?**. The income exempt from garnishment in default $\bar{y}$ is set to the current amount that is exempt in the US, which is 15.1 of GDP per

Table 6: Parameters calibrated jointly in equilibrium

Parameter		**Value**	**Target**	**Data**	**Model**
$\hat{p}$	Minimum expected graduation probability	0.962	Expected graduation rate	*0.860*	*0.860*
$p(e_1)$		0.820		*-0.245*	*-0.244*
$p(e_2)$		0.744		*-0.176*	*-0.176*
$p(e_3)$	Graduation probability parameter | HS GPA	0.828	Sophomores | Enrollment, Table 10	*-0.096*	*-0.095*
$p(e_4)$		0.893		*-0.031*	*-0.031*
$p(e_5)$		0.926		*0.927*	*0.928*
$\rho_d(e_1)$		1.000		*0.570*	*0.672*
$\rho_d(e_2)$		0.696		*0.720*	*0.720*
$\rho_d(e_3)$	Dropout probability persistence | HS GPA	0.657	Seniors / Sophomores by HS GPA	*0.820*	*0.821*
$\rho_d(e_4)$		0.723		*0.870*	*0.871*
$\rho_d(e_5)$		0.705		*0.910*	*0.912*
$\lambda(e_1)$		6.959		*0.090*	*0.080*
$\lambda(e_2)$		5.658		*0.240*	*0.240*
$\lambda(e_3)$	College effort | HS GPA	5.160	College enrollment | HS GPA	*0.440*	*0.441*
$\lambda(e_4)$		4.633		*0.600*	*0.603*
$\lambda(e_5)$		4.134		*0.800*	*0.825*
Z	Aggregate labor productivity	0.303	GDP pc 18 plus	*1.000*	*0.999*
β	Discount factor	0.982	Capital to output	*3.000*	*2.998*
κ	Annual tuition	0.152	Avg. net tuition & fees / GDP pc for 18 plus	*0.088*	*0.088*
$\bar{c}$	College RB consumption	0.147	College RB to GDP pc 18 plus	*0.147*	*0.147*
$\bar{y}$	Exempt income in SL default	0.151	Exempt earnings/GDP pc 18 plus	*0.151*	*0.150*
ϕ	Student loan penalty rate	0.106	Cohort default rate	0.101	0.111
β_c	Altruism for children	0.116	Share of parents' contribution to tuition	*0.335*	*0.336*
σ_ν^2	Variance permanent earnings	0.100	Variance log earnings 25 year olds	0.400	0.390
z	Earnings premium for college	1.452	College wage premium	*1.800*	*1.776*
χ	Social Security replacement rate	0.165	SS expenditure / GDP	*0.041*	*0.041*

capita for the population 18 and over (see Yannelis (2020) for an exact calculation). The parameter that determines a parent's altruism for their child when the child moves out β_c is calibrated to match the share of parents contribution to tuition of 33.5 percent.[11] The variance for the permanent productivity process σ_ν^2 is calibrated such that the model matches the log variance of earnings for 25 year old consumers. The earnings premium for college z is calibrated such that the model matches a college wage premium of 80 percent. The social security replacement rate χ is calibrated such that the model matches total social security expenditure to GDP.

5 Results

In this section, we answer our research question: what is the optimal access to federal student loans in the United States? The answer to this question will take the form of an optimal loan limit, where the limit applies to all consumers but the welfare changes it causes are evaluated from the perspective of each initial skill quintile. First, in section 5.1, we provide a visual illustration of college graduation probabilities and the extent of over-optimism in our model. Second, in section 5.2, we validate our model against untargeted

[11] Source: HSLS, author's calculations; share of students enrolled at a 4-year program whose parents contribute at least sometimes to expenses.

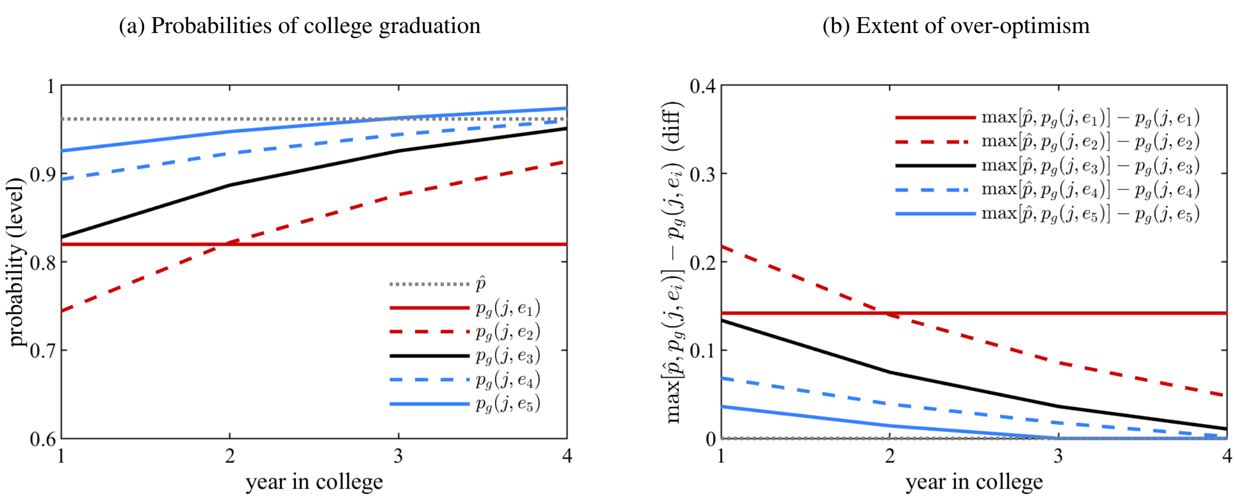
moments related to student loans, and highlight the importance of over-optimism in rationalizing these moments. Third, in section 5.3, we solve for the optimal access to federal student loans by computing transition paths where we vary the student loan limit $\bar{A}$. Finally, in section 5.4, we analyze a transition to the optimal access to student loans.

5.1 Illustration of college graduation probabilities and over-optimism

Figure 1 presents a visual illustration of college graduation probabilities and the extent of over-optimism by HS GPA quintile. Figure 1a plots the true probabilities of graduation $p_g(j, e)$ along with the minimum irrational probability of graduation $\hat{p}$. The true probability of graduation is higher among higher HS GPA quintiles and the conditional true probability of graduation increases with each year of college. Furthermore, the minimum irrational probability of graduation $\hat{p}$ is higher than the true graduation probabilities except for students in the highest HS GPA quintile in their final years of college.

Figure 1b plots the extent of over-optimism in the model, which is computed as the difference between the perceived probability of graduation and the true probability of graduation. Consistent with findings in Stinebrickner and Stinebrickner (2012), our model predicts that students in lower HS GPA quintiles are more over-optimistic. Furthermore, students are more over-optimistic in their early years of college. In the next subsection, we validate the model against non-targeted moments on student loans and highlight the importance of over-optimism in rationalizing key empirical patterns.

Figure 1: Illustration of college graduation probabilities and over-optimism



Notes: Figure 1a plots the true probabilities of graduation $p_g(j, e)$ along with the minimum irrational probability of graduation $\hat{p}$. Figure 1b plots the extent of over-optimism, computed by subtracting the true probability of graduation from the maximum of the irrational probability of graduation and the true probability of graduation.

5.2 Model validation

Table 7 validates the baseline model against untargeted moments related to student loans in the data. It also reports the same statistics for the baseline model without over-optimism ($\hat{p} = 0$). The first set of moments, from the HSLS, report the share of fourth year students in college with positive federal student loan balances (debtors) and the average outstanding balance, conditional on a positive balance, by HS GPA quintiles. These moments inform us about the utilization rates of federal student loans. The model does fairly well in rationalizing the share of fourth year students with federal student loans in the first two HS GPA quintiles (74.00-84.00 percent in the data and 88.32-93.22 percent in the model). However, the model overstates the share of debtors in the third to fifth HS GPA quintiles (54.00 to 65.00 percent in the data and 91.63-92.79 percent in the model). While overstating the share of debtors in higher HS GPA quintiles, the model does remarkably well in rationalizing the intensive margin of student loan balances across all HS GPA quintiles. The average balance for a fourth year student with a positive balance is roughly 20 percent of GDP per capita across HS GPA quintiles in both model and data.[12]

The second set of moments, from the SCF, report the distribution of debtors and student loan balances by dropouts, college graduates, and those not in repayment (i.e., enrolled in college). Again, the model does reasonably well in rationalizing these moments given that they were not targeted in our calibration. The share of dropouts among those with a positive student loan balance is 18.35 percent in the data and 11.44 percent in the model. The share of outstanding student loan balances held by dropouts is 10.79 percent in the data and 8.19 percent in the model. The third set of moments, also from the SCF, report the share of student loan defaults by dropouts and college graduates (those enrolled in college are not relevant for this statistic). The model rationalizes roughly half of the share of defaults by dropouts observed in the data (51.27-57.72 percent in the data and 24.23 percent in the model). Hence, even with over-optimism, the model somewhat understates the shares of student loans and defaults accounted for by dropouts.

A natural question to ask is to what extent is the over-optimism in our model important in accounting for these moments. Therefore, in the last column of Table 7, we report the same moments when we solve the baseline model without over-optimism ($\hat{p} = 0$) and keep all other parameters the same. Most interestingly, the share of dropouts among those with a positive student loan balance decreases from 11.44 to 8.91 percent. The share of outstanding student loan balances held by dropouts decreases from 8.19 to 6.82 percent. And finally, the share of defaults by dropouts decreases from 24.23 to 15.18 percent. Hence, in our calibration, over-optimism is important in rationalizing student loans and defaults among dropouts.[13]

5.3 Solving for the optimal access to federal student loans

Having highlighted the significance of over-optimism in rationalizing student loan balances and defaults among dropouts in the previous section, we now characterize the optimal level of access to federal student

[12] If we incorporate college room and board as a mandatory expense or if the student loan limit is equal to that of an independent student, as in Chatterjee and Ionescu (2012) and Krueger and Ludwig (2016), the model significantly overstates the utilization of federal student loans. Hence, the use of HSLS data allows us to discipline our modeling choices along these dimensions.

[13] In the model without over-optimism, none of the students in the lowest HS GPA quintile enroll in college. Hence, students loans are not held by those in the lowest HS GPA quintile.

Table 7: Untargeted moments

Variable (unit = percentage)	**Data**	**Baseline model**	**No over-optimism model**
Panel A. Fourth-year college students debtor status and loan balance			
With positive federal SL ∣ HS GPA quintile 1	84.00	93.22	n/a
With positive federal SL ∣ HS GPA quintile 2	74.00	88.35	71.83
With positive federal SL ∣ HS GPA quintile 3	65.00	91.63	85.56
With positive federal SL ∣ HS GPA quintile 4	65.00	92.80	92.71
With positive federal SL ∣ HS GPA quintile 5	54.00	92.79	94.55
Avg. federal SL ∣ [HS GPA quintile 1 & SL > 0] / GDP pc 18 plus	18.80	26.47	n/a
Avg. federal SL ∣ [HS GPA quintile 2 & SL > 0] / GDP pc 18 plus	19.82	20.76	19.25
Avg. federal SL ∣ [HS GPA quintile 3 & SL > 0] / GDP pc 18 plus	22.51	23.27	23.90
Avg. federal SL ∣ [HS GPA quintile 4 & SL > 0] / GDP pc 18 plus	22.61	24.01	26.15
Avg. federal SL ∣ [HS GPA quintile 5 & SL > 0] / GDP pc 18 plus	21.19	22.94	26.16
Panel B. Educational attainment in the population of debtors			
Share of HH: dropouts	18.35	11.44	8.91
Share of HH: enrolled or in grace period	16.35	16.26	17.75
Share of HH: graduated	65.30	72.29	73.34
Share of $: dropouts	10.79	8.19	6.82
Share of $: enrolled or in grace period	15.77	12.22	13.32
Share of $: graduated	73.44	79.58	79.86
Panel C. Educational attainment in the population of delinquent debtors			
Dropouts	51.27-57.72	24.23	15.18
Graduates	42.28-48.73	75.77	84.82

Panel A presents information on debt accumulation and loan balance at the college graduation age. Panel B presents information on the composition of all households with a positive balance of educational debt, and how the aggregate balance of student debt is distributed across different types of households. Panel B presents statistics for households that are not making payments on their educational loans because they cannot afford to do so (delinquent households).

loans. Our measure of access is the cumulative student loan limit, $\bar{A}$. Apart from general equilibrium effects,[14] an expansion in student loan limits might be suboptimal in our framework for two reasons. First, students are over-optimistic about graduation probabilities. Second, the student loan interest rate is set by the government and does not reflect default risk. To solve for the optimal access in the presence of these effects, we analyze transitions paths after to new steady states various changes in the student loan limit. We assume that the economy is in initial steady state in $t = 0$. The transition is unexpectedly announced in $t = 1$, but agents have perfect foresight thereafter. We then compute the welfare changes of the average 18 year-old consumer within each HS GPA quintile.

Our measure of welfare is consumption equivalent variation. It measures the lifetime change in consumption required in the initial steady state in every period and every state for a consumer to be indifferent between

[14] In this model, general equilibrium effects operate via prices, taxes, accidental bequests, and social security benefits.

the initial steady state and the transition (or the final steady state).[15] Importantly, when measuring welfare for the 18-year-old consumer, we assume that the government is paternalistic. That is, the government internalizes that the consumer is overly optimistic when making the college enrollment decision, but knows what the true payoff of those choices is. The government will use the policy function $\hat{d}_s$ that solves the over-optimistic consumer's college entrance problem (equation 6) but the realized utility will be based on the value function with true graduation probabilities. Thus, from the government's perspective, the value to a 18-year-old is given by

$$W(\nu,\eta,a,e) = (1-\hat{d}_s)V(1,\ell,\nu,\eta,a) + \hat{d}_s V(1,h,\nu,\eta,a,e), \tag{19}$$

where $\hat{d}_s$ is the policy function that sovles equation 6 and $V(1,h,\nu,\eta,a,e)$ is the value of college from using the true probabilities of graduation in equation 9.

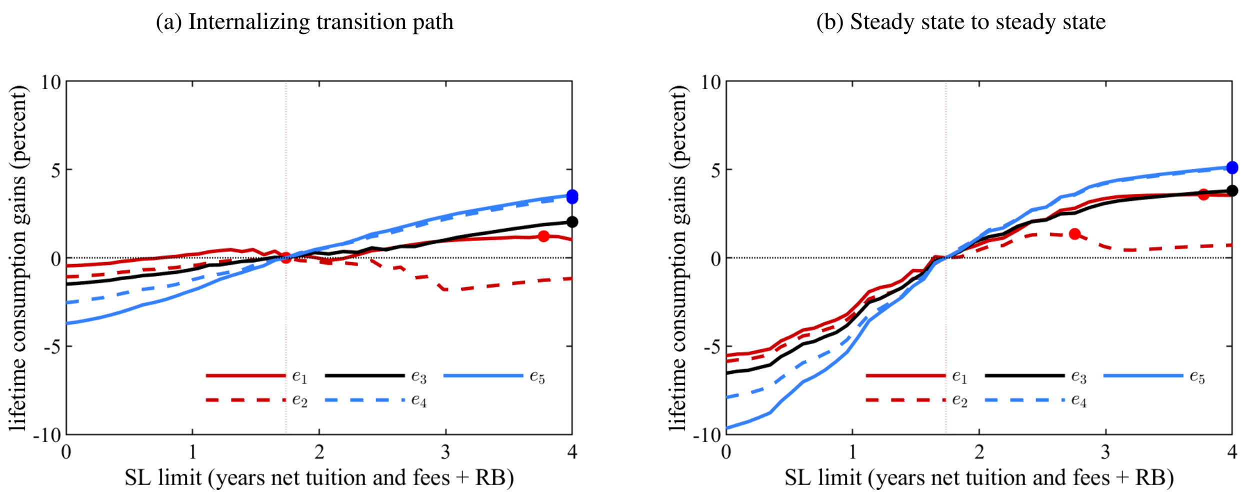
Figure 2 plots the welfare implications to an average 18-year-old by HS GPA quintiles from our policy experiments. The left panel, Figure 2a, plots the welfare implications when the 18-year-old internalizes the transition path and the right panel, Figure 2b, plots the welfare implications computed by comparing the initial steady state to the final steady state. The former serves as a measure of welfare for 18-year-old consumers that undergo the transition to the new student loan limit and the latter serves as a measure of welfare for future generations. In both panels, the x-axis is the new student loan limit in units of the number of years of tuition and fees plus room and board. For example, in the baseline model, the student loan limit is 1.737 implying the limit can pay for at most 1.737 years of tuition and fees plus room and board. We analyze transitions to student loan limits that vary from 0-4 years tuition and fees plus room and board (i.e., eliminating the federal student loan program to the federal student loan program being sufficient to fully pay for college). In both panels, each line represents lifetime consumption gains/losses to the average 18-year-old consumer by HS GPA quintiles.

Both Figures 2a and 2b show that expanding limits leads to gains and tightening limits leads to losses monotonically for consumers in the top three HS GPA quintiles. Consumers in the lowest HS GPA quintile benefit from increasing the student loan limit to almost 4 years of tuition. Consumers in the second HS GPA quintile prefer the status quo limit when internalizing the transition path and a marginal expansion in limits when comparing steady states. Hence, the main takeaway from these figures is that, even with over-optimism and student loan interest rates that do not reflect default risk, for the average 18-year-old consumer, it is optimal to increase student loan limits to be enough to fully pay for college, i.e., the optimal limit is 4 years of tuition and fees plus room and board in our model.

Another takeaway from Figure 2 relates to the magnitude of gains/losses when internalizing transition paths versus comparing steady states. The gains from expanding the limit to 4 years of college when internalizing the transition path range from -1.2-3.5 percent of lifetime consumption, whereas when comparing steady states, the gains range from 0.7-5.1 percent of lifetime consumption. Therefore, the gains from expanding student loan limits to the optimal one are larger in the long run. The economic intuition for this result is

[15] To be precise, as our utility function includes an effor cost for college $\lambda(e)$, if consumption increases by g, it also implies the effort cost scales by $(1+g)^{1-\sigma}$. With $\sigma > 1$, this implies that if consumption increases by g, the effort cost is divided by $(1+g)^{\sigma-1}$.

Figure 2: Welfare implications of changing the federal student loan limit, average 18-year-old, by HS GPA quintiles in baseline model



Notes: Figure 2 plots the welfare implications to an average 18-year-old given their HS GPA quintile from our policy experiments of transitioning to a new student loan limit in the baseline model. The left panel plots the welfare implications when the transition path is internalized and the right pane plots the welfare implications when the transition path is not internalized. The x-axes are the new student loan limit in units of the number of years of tuition and fees plus room and board. The y-axes are lifetime consumption gains/losses to the average 18-year-old consumer by HS GPA quintiles. The round marker on each line indicates the student loan limit that maximizes welfare for the respective skill quintile.

discussed in Section 5.4.

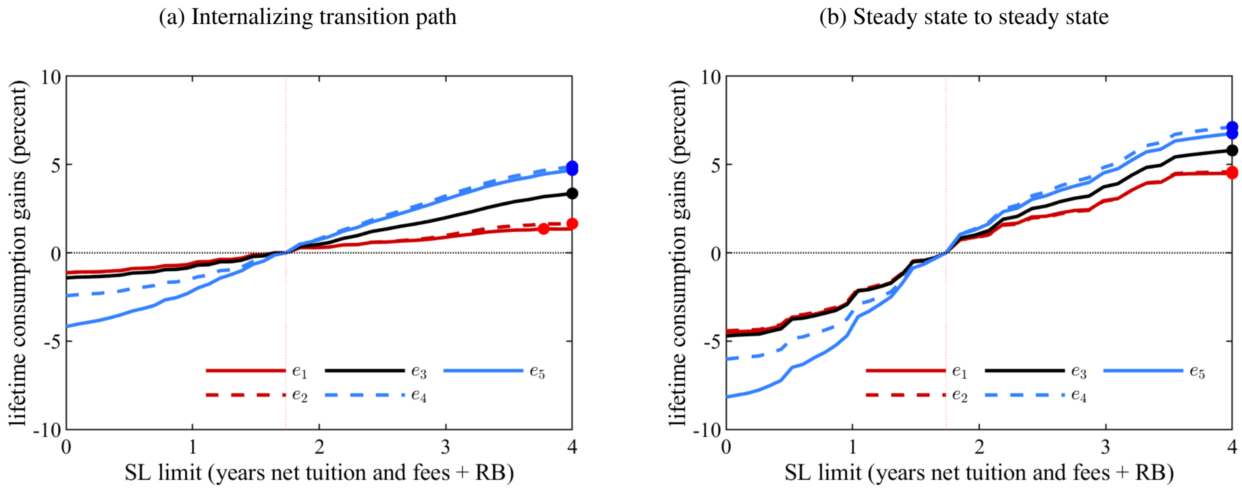
Figure 3 plots the welfare implications analogous to those of Figure 2 but when there is no over-optimism ($\hat{p} = 0$) and all other parameters are the same. In this case as well, the optimal access is expanding the student loan limit to 4 years of college. This result, while not surprising, is not obvious because the student loan interest rate is set by the government and does not reflect default risk.

5.4 Transition to the optimal access to federal student loans

In this section, we analyze the transition to the optimal access to federal student loans. Based on the previous section, the optimal limit is 4 years of tuition and fees plus room and board for most 18-year-old consumers. Hence, we analyze a transition from the status quo to a student loan limit where $\bar{A} = 4$.

Figure 4a plots the lifetime consumption gains to the average 18-year-old along the transition path. The consumption gains increase from 1.8 percent in the period of the transition to almost 4 percent in the long run. Figure 4b, which plots the average income tax rate for a consumer with income equal to gdp per capita and the prices (interest rate and wage rate), provides economic intuition for why the gains increase in the long run. The income tax rate, initially, increases marginally to finance the increase in public subsidies for education due to higher enrollment. In the long run, the income tax rate decreases by almost 2 percentage points. This leads to larger gains in the long run. The reason for the fall in income taxes is the increase in the tax base due to an increase in college graduates (i.e., higher earnings, savings, and consumption lead to

Figure 3: Welfare implications of changing the federal student loan limit, for average 18-year-old, by HS GPA quintiles in no over-optimism model ($\hat{p} = 0$)



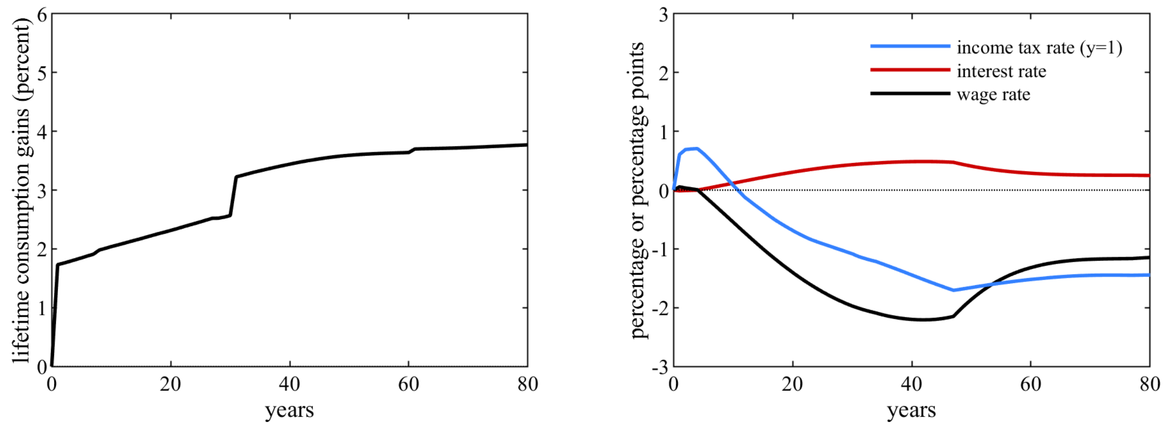
Notes: Figure 3 plots the welfare implications to an average 18-year-old given their HS GPA quintile from our policy experiments of transitioning to a new student loan limit in the model without over-optimism (baseline model with $\hat{p} = 0$). The left panel plots the welfare implications when the transition path is internalized and the right panel plots the welfare implications when the transition path is not internalized. The x-axes are the new student loan limit in units of the number of years of tuition and fees plus room and board. The y-axes are lifetime consumption gains/losses to the average 18-year-old consumer by HS GPA quintiles. The round marker on each line indicates the student loan limit that maximizes welfare for the respective skill quintile.

higher income and consumption tax revenues). The wage rate, which falls in the long run due to an increase in effective labor supply because of an increase in the number of college graduates, could only dampen the welfare gains in the long run. The interest rate on savings increases in the long run, which generates small additional welfare gains.

Figure 4: Transition path to optimal access to federal student loans

(a) Welfare implications, average 18-year-old

(b) Income tax rate and prices



Notes: Figure 4a plots the lifetime consumption gains to the average 18-year-old along the transition path to the optimal student loan limit (4 years of college and fees plus room and board). For the same transition, Figure 4b plots the average income tax rate for a consumer with income equal to gdp per capita and the prices (interest rate and wage rate).

6 Conclusion

In this paper, we examine the optimal student loan limit in an environment where low-skill students are overly optimistic about their probability of finishing college. This over-optimism means that they are more likely to be repaying their student loan debt without enjoying the college wage premium. As a consequence, when given the opportunity, some students over-borrow, which means that low-skill students tend to benefit less from loan limit expansions than their higher-skilled peers.

Here, we focus on the student loan limit for our policy analysis, but we could examine other margins—for example, student loan forgiveness, penalties for non-repayment, or interest rates that incorporate non-completion risk. Focusing on the optimal use of other policy instruments would potentially highlight different consequences for over-optimism among young adults when they make the college enrollment decision.

References

Athreya, K., C. Herrington, F. Ionescu, and U. Neelakantan (2021). Student loan borrowing and repayment decisions: Risks and contingencies. *The Handbook of Economics of Education, Edited by Brian McCall*.

Caucutt, E. and L. Lochner (2020). Early and Late Human Capital Investments, Borrowing Constraints, and the Family. *Journal of Political Economy 128*(3), 1065–1147.

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

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics 129*(4), 1553–1623.

Conesa, J. C., D. Costa, P. Kamali, T. J. Kehoe, V. M. Nygaard, G. Raveendranathan, and A. Saxena (2018). Macroeconomic effects of medicare. *Journal of Economics of Ageing 11*, 27–40.

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.

Exler, F., I. Livshits, J. MacGee, and M. Tertilt (2021). "consumer credit with over-optimistic borrowers. *Unpublished manuscript*.

Friends of Bernie Sanders (2021). Issues: College for All and Cancel All Student Debt. Accessed: 12-30-2021.

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.

Ionescu, F. (2009). The federal student loan program: Quantitative implications for college enrollment and default rates. *Review of Economic Dynamics 12(1)*, 205–231.

Ionescu, F. (2011). Risky human capital and alternative bankruptcy regimes for student loans. *Journal of Human Capital 5(2)*, 153–206.

Ionescu, F. and N. Simpson (2016). Default risk and private student loans: Implications for higher education policies. *Journal of Economic Dynamics and Control 64*, 119–147.

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.

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

Luo, M. and S. Mongey (2019). On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium. pp. Unpublished manuscript.

Nakajima, M. (2012). Rising indebtedness and temptation: A welfare analysis. *Quantitative Economics 3(2)*, 257–288.

Nakajima, M. (2017). Assessing bankruptcy reform in a model with temptation and equilibrium default. *Journal of Public Economics 145*, 42–64.

Schumer, C. and E. Warren (2020). *Calling on the President of the United States to take executive action to broadly cancel Federal student loan debt.*

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).

Stinebrickner, T. and R. Stinebrickner (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics 30(4)*, 707–748.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters 20(2)*, 177–181.

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics (2019). Table 330.10. Accessed: 1-4-2022.

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics (2020). Table 331.35. Accessed: 1-4-2022.

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

A Data Appendix

A.1 HSLS

The High School Longitudinal Study of 2009 (HSLS) is a representative panel of 9th-grade students in the United States beginning in 2009. The survey collection occurs over several waves, with the most recent wave being the collection of post-secondary transcripts in the 2017-2018 academic year. If sample members begin a 4-year degree program in the fall after high school graduation (the fall of 2013), and does not take time off from school, then this transcript collection occurs during the fall of their fourth and final year of college. Table 8 contains an outline of the structure of the HSLS.[16]

As the outline indicates, the focal sample member is referred to as "Student" while they are in high school and as "Sample member" during the 2013 update, because they are between educational programs. Regardless of the focal individual's educational status after the base year, the HSLS makes an effort to collect data from them. Thus the second follow-up in 2016 includes information from students who are currently enrolled in post-secondary education as well as those who are not.

Table 8: Structure of the HSLS

Calendar Year(s) **Academic Year (if enrolled)** **Questionnaire**	**Base Year** **2009 (Fall)** **1st year HS (Fall)**	**1^{st} Follow-up** **2012 (Spring)** **3rd year HS (Spring)**	**2013 Update** **2013 (Summer)** **Graduated HS**	**HS Transcripts** **2013-2014** **Graduated HS**	**2^{nd} Follow-up** **2016** **3rd year PS (Spring)**	**PS Transcripts** **2017-2018** **4th year PS (Fall)**
Student	X	X				
Parent	X	X				
Student/Parent			X			
Sample Member					X	
Counselor [1]	X	X				
Aministrator [2]	X	X				
Teacher [3]	X					
Institution Attended				X		X

Notes: Table 8 describes the survey structure of the HSLS. Here, HS stands for High School and PS stands for Post-secondary. Calendar years and academic years are distinguished because academic years overlap two calendar years and the semester of data collection is indicated for the academic year except for the 2013 Update and HS transcript collection when this is not relevant. The focal sample member is referred to as "Student" while they are in high school and as "Sample member" during the 2013 update, because they are between educational programs. Regardless of the focal individual's educational status after the base year, the HSLS makes an effort to collect data from them. This ensures that the second follow-up in 2016 includes information from students who are currently enrolled in post-secondary education as well as those who are not.

[1] Lead counselor at student's high school.
[2] Administrator (principal) at student's high school.
[3] Math or science teacher at student's high school.

A.2 SCF

Table 9 contains variable codes corresponding to conceptual categories discussed in the text. Each individual loan is assigned a flag (non-completion, non-repayment) and then the portfolio of loans is assigned to a category (non-completion, non-repayment) according to a threshold rule. The preferred specification we

[16] Questionnaires are available here: https://nces.ed.gov/surveys/hsls09/questionnaires.asp.

use for this threshold rule is described in the text.

Table 9: 2019 Survey of Consumer Finances: Variables used

Variable	Description
Survey	
Weight	X42001
Income + Components	
Total HH Income (manual)	(A)+(B)+(C)+(D)
Wages and Salaries (A)	X5702
Self Emp. Income (B)	X5704
Capital Income (C)	X5706 + X5708 + X5710 + X5712 + X5714
Other (D)	X5724 + X6558 + X6566 + X6574 + max(0,X6464) + max(0,X6469)+ max(0,X6474) + max(0,X6479) + max(0,X6965) + max(0,X6971) + max(0,X6977) + max(0,X6983)
Wealth: Education Loans	
Current balance	X7824 +X7847 +X7870 +X7924 +X7947 +X7970 +X7179
Loan for: $\mathbb{I}_{\text{resp,sp}}$	if=1 or 2: X7978, X7883, X7888, X7893, X7898, X7993
Flag fed loan	if=1: X7879, X7884, X7889, X7894, X7899, X7994
Flag late pymt can't afford	if =3: X7821, X7844, X7867, X7921, X7944 , X7967
Flag for did not complete program loan was taken out to pay for	if =5: X7881, X7886, X7891, X7896, X7901, X7996

B Model Appendix

B.1 Value functions

The over-optimistic value of college for $j = 4$ is given by

$$\hat{V}(j,h,\nu,\eta,a,e) = \max_{\hat{c},\hat{a}'} U(c,j,h) + \beta\psi_j\big[\max[p_g(j,e),\hat{p}]E_{\eta'|h,\eta}V(j+1,h,\nu,\eta',\hat{a}',e) \tag{20}$$
$$+ (1-\max[p_g(j,e),\hat{p}])E_{\eta'|\ell,\eta}V(j+1,\ell,\nu,\eta',\hat{a}')\big]$$
$$s.t.$$
$$(1+\tau_c)\hat{c} + \hat{a}' + (1-\theta(e)-\theta^{pr}(e))\kappa = y_{j,h,\nu,\eta,a} + a + Tr_j - T(y_{j,h,\nu,\eta,a})$$
$$\hat{a}' \geq -\frac{\bar{A}[(1-\theta(e)-\theta^{pr}(e))\kappa + \bar{c}]j}{4}$$
$$\hat{c} \geq 0.$$

The rational value of college for $j = 4$ is given by

$$V(j,h,\nu,\eta,a,e) = \max_{c,a'} U(c,j,h) + \beta\psi_j\big[p_g(j,e)E_{\eta'|h,\eta}V(j+1,h,\nu,\eta',a',e) \tag{21}$$
$$+ (1-p_g(j,e))E_{\eta'|\ell,\eta}V(j+1,\ell,\nu,\eta',a')\big]$$
$$s.t.$$
$$(1+\tau_c)c + a' + (1-\theta(e)-\theta^{pr}(e))\kappa = y_{j,h,\nu,\eta,a} + a + Tr_j - T(y_{j,h,\nu,\eta,a})$$
$$a' \geq -\frac{\bar{A}[(1-\theta(e)-\theta^{pr}(e))\kappa + \bar{c}]j}{4}$$
$$c \geq 0.$$

The value of default for $j = j_f + j_a$ is given by

$$V^D(j,s,\nu,\eta,a,e) = U(c,j,s) + \beta\psi_j E_{\eta'|s,\eta}V(j+1,s,\nu,\eta',a') \tag{22}$$
$$+ \beta_c E_{\nu'|\nu}E_{\eta'|l}W(\nu',\eta',b,e) \tag{23}$$
$$s.t.$$
$$(1+\tau_c)c = y_{j,s,\nu,\eta,a} + Tr_j - T(y_{j,s,\nu,\eta,a}) - \rho_D(a, y_{j,h,\nu,\eta,a})$$
$$a' = (1+r_{SL})a + \rho_D(a, y_{j,h,\nu,\eta,a}) - \phi\max[\rho_R(j,a) - \rho_D(a, y_{j,h,\nu,\eta,a}), 0]$$
$$b = 0.$$

B.2 Equilibrium definition

To define the equilibrium, we must first discuss more notation and define the social security function. Let $\overrightarrow{\omega}$ denote the idiosyncratic state of an agent. This state depends on age and enrollment status in the following way:

$$
\overrightarrow{\omega} = \begin{cases} (j,\nu,\eta,a,e) & \text{for 18-year-olds, before making the college entrance decision} \\ (j,h,\nu,\eta,a,e) & \text{for agents in college} \\ (j,s,\nu,\eta,a) & \text{for agents not enrolled, dropouts, or graduates, unless } j = j_f + j_a \\ (j,s,\nu,\eta,a,e) & \text{if } j = j_f + j_a \end{cases} \tag{24}
$$

The social security function is given by

$$
ss_{s,\nu} = \chi \Bigg[\frac{.5 \int w\nu\eta\epsilon_{j,s}[1_{s=h}z + 1_{s=\ell}]\Omega_t d(\overrightarrow{\omega}|18 \leq j < j_r, s, \nu)}{\int \Omega_t d(\overrightarrow{\omega}|18 \leq j < j_r, s, \nu)} + \tag{25}
$$

$$
\frac{.5 \int w\nu\eta\epsilon_{j,s}[1_{s=h}z + 1_{s=\ell}]\Omega_t d(\overrightarrow{\omega}|18 \leq j < j_r)}{\int \Omega_t d(\overrightarrow{\omega}|18 \leq j < j_r)} \Bigg],
$$

where it replaces χ fraction of the average of the average labor earnings for the 30 years before retirement conditional on education and permanent productivity and the average unconditional labor earnings for the 30 years before retirement.

Definition Given an initial level of capital stock K_0 and initial distribution over idiosyncratic states $\Omega_0(\overrightarrow{\omega})$, a competitive equilibrium consists of sequences of household value functions $\{W_t(\overrightarrow{\omega}), V_t(\overrightarrow{\omega}), \hat{V}_t(\overrightarrow{\omega}), V_t^R(\overrightarrow{\omega}), V_t^D(\overrightarrow{\omega})\}$, household college entrance and dropout policy functions $\{\hat{d}_{s,t}(\overrightarrow{\omega}), \hat{d}_{d,t}(\overrightarrow{\omega}), d_{d,t}(\overrightarrow{\omega})\}$, household consumption and next period asset policy functions $\{\hat{c}_t(\overrightarrow{\omega}), \hat{a}'_t(\overrightarrow{\omega}), c_t(\overrightarrow{\omega}), a'_t(\overrightarrow{\omega})\}$, household default policy function $\{d_{f,t}(\overrightarrow{\omega})\}$, household inter vivos transfer policy function $\{b_t(\overrightarrow{\omega})\}$, production plans $\{Y_t, K_t, L_t\}$, sequence of federal student loan policies $\{\bar{A}_t, r_{SL}, \bar{y}, \tau_g, T_{SL}\}$, sequence of tax, government consumption, public education subsidy, and social security policies $\{\tau_p, \gamma_t, \tau_c, g, \theta(e), \chi\}$, sequence of prices $\{r_t, w_t\}$, sequence of accidental bequests $\{Tr_{t,j}\}$, and sequence of 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 6, 7, 8, 9, 10, 11, 12, 13, 14, 20, 21, 22;

(ii) Interest rate and wage rate satisfy equations 16 and 17, respectively;

(iii) Accidental bequests are transferred to households between ages 50 and 60 ($33 \leq j \leq 43$) after deducting expenditure on private education subsidies:[17]

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

where $N_{t,j}$ denotes the mass of population at time t of age j;

[17] In our baseline calibration, and 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, then bequests would be negative, which is equivalent to a lump-sum tax.

(iv) Government budget constraint balances as follows by adjusting γ:

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

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, and are computed as follows:

$$\begin{aligned}
G_t =& gY_t = gK_t^{\alpha}(zL_t)^{1-\alpha} \\
E_t =& \kappa \int \theta(e) 1_{s=h \text{ and } j\in\{1,2,3,4\}} \Omega_t d(\overrightarrow{\omega}) \\
D_t =& \int 1_{j\leq 4}\Big[[\min[a,0] - \min[a_t'(\overrightarrow{\omega}),0]] + \\
& 1_{j>4}(1-d_{f,t}(\overrightarrow{\omega}))[\min[a,0](1+r_{SL}) - \min[a_t'(\overrightarrow{\omega}),0]] + \\
& 1_{j>4} d_{f,t}(\overrightarrow{\omega})[-\rho_D(a, y_{t,j,s,\nu,\eta,a}) + \phi \max[\rho_R(j,a) - \rho_D(a, y_{t,j,s,\nu,\eta,a}),0]\Big]\Omega_t d(\overrightarrow{\omega}) \\
SS_t =& \int 1_{j\geq j_r} ss_{t,s,\nu}\Omega_t d(\overrightarrow{\omega});
\end{aligned}$$

(v) Labor, capital, and goods markets clears in every period t:

$$\begin{aligned}
L_t =& \int [1_{j\leq 4, s=h}\nu\eta\epsilon_{j,\ell} l_{pt} + 1_{4<j<j_r, s=h}\nu\eta\epsilon_{j,s} z + 1_{j<j_r, s=\ell}\nu\eta\epsilon_{j,s}]\Omega_t d(\overrightarrow{\omega}) \\
K_{t+1} =& \int a_t'(\overrightarrow{\omega})\Omega_t d(\overrightarrow{\omega}) \\
Y_t =& C_t + K_{t+1} - (1-\delta)K_t + G_t + \kappa \int 1_{j\leq 4, s=h}\Omega_t d(\overrightarrow{\omega}) + \\
& \phi \int 1_{j>4} d_{f,t}(\overrightarrow{\omega}) \max[\rho_R(j,a) - \rho_D(a, y_{t,j,s,\nu,\eta,a}),0]\Omega_t d(\overrightarrow{\omega}),
\end{aligned}$$

where C_t is aggregate consumption; and

(vi) $\Omega_{t+1} = \Pi_t\Big(\Omega_t\Big)$, where Π_t is the law of motion that is consistent with the household policy functions and the exogenous processes for population, labor productivities, HS GPA, and college dropouts.

C Calibration Appendix

C.1 Regressions for indirect inference

Table 10 reports how persistence to year 2 of college conditional on year 1 enrollment is predicted by student skill quintile (with the highest quintile dropped), family income, parent education, and (for model 2) hours worked and employment status.

Table 10: Enrolled in Year 2 | Enrolled in Year 1 (With and without hours)

Variable	**(1) With Hours**	**(2) Without Hours**
HS GPA Quintile = 1	-0.228***	-0.245***
	(0.0546)	(0.0566)
HS GPA Quintile = 2	-0.166***	-0.176***
	(0.0349)	(0.0344)
HS GPA Quintile = 3	-0.0887***	-0.0956***
	(0.0200)	(0.0197)
HS GPA Quintile = 4	-0.0272**	-0.0311**
	(0.0136)	(0.0135)
Average family income during HS	2.10e-08	1.93e-08
	(1.60e-08)	(1.59e-08)
At least 1 parent AA+	0.00577	0.0106
	(0.0292)	(0.0298)
Average hours worked per week in current year	-0.00272***	
	(0.00105)	
Not working in current year	-0.0683***	
	(0.0213)	
Constant	0.988***	0.927***
	(0.0326)	(0.0297)
r^2	0.05	0.04
Observations (rounded)	4430	4430

Notes: Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study Restricted-Use Data File.