

Groupe 37

MONTEIRO Lucas

CHRIFI ALAOUI Mohamed-Amin

AZEVEDO Noah

PINAULT Raphaël

Impact of the student loan on the labor income

Abstract

This study is dedicated to the analysis of the impact of the student loan on the labor income for individuals between 25 and 35 years old. The dataset is provided by the Panel Study of Income Dynamics (PSID). We record answers from 2427 families in the United States in 2017. Using robust OLS, after controlling for the level of education, the major, the education of the individuals' parents, the sex and age of the individual, we found that *ceteris paribus* an increase of the student loan of 1 percent increases the labor income of 0,01 percent. Our approach is not completely free from the endogenous problems.

Table of content

I.Introduction	3
II. Contribution of the literature	4
III. Presentation of the data	6
IV. Econometric models and results	7
V. Conclusion and limits	9
VI. Tables and Figures	11
VII. References	13

I.Introduction

Contracting a student loan is a mandatory step for many people in the United States, given the important cost of college education. Since the 90s we noticed a rise of 16% of the number of undergraduate students that took a loan. On one hand, students that contract a loan have access to better education awarding them with more thought after, higher-yielding skills and easier insertion into the labor market. Thus, we could imagine a positive relation between student loans and income. On the other hand, by contracting a loan the individual will be constrained by the reimbursement for many years after graduation. Leading to precarity implying a negative effect of the loans on income [source]. The importance of this decision is amplified by the fact that the decision is non-reversible.

This ambiguity is already noticeable in the range of the research done in the past. Students with student debts would go on to study for higher paying jobs and avoid low-paying jobs ([Rothstein, Rouse \(2011\)](#)) let us think of a positive relationship between student loans and labor income, while the fact that young adults with student loans may have a lesser net worth than those with no student loans ([Elliott, Nam \(2013\)](#)) suggests a negative relationship between student loans and labor income.

Therefore, the effect of student loans on future labor income can be ambiguous. The interest of this paper is to analyze whether the final effect of student loans on income is positive or negative. This analysis can assist the students to make the correct decisions regarding the student loans and their future income. In addition, it can also advise the lenders, generally a privat bank, in the decision of providing the loan.

Our study is articulated around three main sections. First, we go through contemporary literature and see how it contributed to the subject. Then, we present our data with a detailed

explanation of the variables that will be used. The third and last section is dedicated to the presentation of our econometric models and the main results. Finally, a conclusion at the end of the study enables us to summarize our main results and to go through the limits of our work.

II. Contribution of the literature

Results from academic research are diverse but not contradictory. On one hand, [Erin et al. \(2019\)](#) show that additional debt is associated with a higher likelihood of having a negative net worth. [Elliott and Nam \(2013\)](#) expand this idea by adding that young adults with student loans may have a lesser net worth than those with no student loans. Being more specific, [Minicozzi \(2005\)](#) shows that an increase amounting to \$1000 of the educational debt (that is the debt associated with student loan) yields to an increase of the labor income directly after leaving school by nearly 1%. It has been shown that taking a student loan would raise the credit scores by 24 points ([Brown, Caldwell \(2013\)](#)). Besides, according to [Abel and Deitz \(2014\)](#), the benefits of college outweigh the costs of the educational debt.

Students take loans in order to access their desired degree and ultimately to be in a better disposition to have the career of their choice ([Britton et al. \(2019\)](#)). The choice of their future career and the amount of the loan is thus taken simultaneously in a certain regard. This has two implications and consequences on the choice of the regression model down the road. The amount of the student loan is decided according to the tuition fee of the desired major degree but also depending on the potential future income perceivable after graduation the person has already in mind ([Rothstein, Rouse \(2011\)](#)). In other words, the student loan is partially explained by the expectation of the person's future labor income as well as the choice of the major degree and level of education. Similarly, after graduation, the individual will seek jobs while being constrained by the student loan credit. In this regard, the labor income depends on the student loan. All in all, this translates into an endogeneity problem that needs to be taken

into account in the specification. Otherwise, we would expect that the impact of the student loan to be positively overestimated.

As explained, the choice of the major and the level of education are major confounders. Which itself reveals to be dependent on the expectation of the income the individual has. In other words, the choice of the major and the level of education is partially made in consideration of the expectation of the perceivable future labor income. This translates into a second endogeneity problem that should be addressed down the road. To a certain extent, incomes are specific to each industry. Industry such as finance or high technology yields much higher salaries compared to other industries such as teaching (Avery et al, 2012). It has been discussed that the choice of the major and at last the industry in which the students will end up after their graduation, is made according to the amount of the student loan; reason being that they need to make sure to be able to reimburse the loan (Rouse et al, (2004)). If we omit the sector and the level of education, we suspect the impact of the student loan on the labor income to be smaller.

Educated parents are aware of these subjects and evidence from Abbott et al. (2019) and Kean and Wolpin (2001) suggest that intergenerational transmission of abilities may be affected by parental wealth and education. The procurement and the amount of the student loan depends on the level of guarantee the student can testify. Subsequent to the fact that, generally, parents are the guarantor, it is clear that the amount of the student loan also depends on the income or more generally the level of wealth of the parents.(Rothstein et al. (2011) ; Houle, Jason (2014)) Which in turn also depends on their respective level of education. On top of that, the choice of the level of education and the major is dependent on the education of the parents. Often, parents with a high level of education require their children to have at least the same level of education (Sirin, (2005) ; Behrman, (1990)). Thus we suspect that the education of the parents has a positive impact on the income through the student loans. By including controls for the parents' education we expect the student loan coefficient to be smaller.

Another confounder that it would be interesting to include in our specification is the sex of the individual. There's an over representation of male students in scientific, finance and technologic majors; which is by far the one that leads to the highest incomes in the industry. Numerous papers reported the income gap due to the sex of the individual ([Weichselbaumer et al. \(2005\)](#)). This gap originates from social inequalities. Thus, the sex of the individual is both a deterministic factor for the income as well as the choice of the major ([Chris Sakellariou et al. \(2021\)](#)).

Additionally, we can imagine that having a child at charge may impose a constraint on the career of the individual since more time will be devoted to the child.

III. Presentation of the data

Our data comes from the Panel Study Income Dynamic (PSID) for the year 2017, in the United States of America. After keeping the variables of interest for our study and removing the individuals with missing information, we went from a data set of 9607 observations to one with 1918 observations, observations that are at the household level, for a total of 7 variables and controls.

Our variable of interest (or dependent variable) is the log of the labor income per head, while our independent variable is the log of the amount of the student loan per head. The use of the logarithm reduces the heteroscedasticity. Moreover, we analyze these variables “per head” in order to have coherent information between the households composed of two parents and the households composed of only one individual. Hence, for the households composed of two individuals, we simply take the average of their income and their student loan.

The variable *Highest Degree* (NoCollege, Bachelor, Master, PhD and other) is a proxy for the highest degree obtained by the referent, while the variable *Major* (Engineering,

Business/Finance and other) measures the field of study chosen by the referent. Moreover, we also control for the education of the parents of the referent with the variable *Parents' Education* (Basic, High School and College). Besides, the variable *Number of Children* enables us to control on how many children are in the household, and the variable *Sex* is a control for the sex of the referent. We also have a variable *Age*. In our analysis, we took an interval between 25 and 35 years.¹ For this, we wanted to focus on the beginning of the individual's career. Certainly, when one wants to look at an individual's future income, one might think that it would be more ideal to look at his or her income over a lifetime to see the long-term effects. But, for us, we think that we can lose the direct impact of the student loan because in the future the experience (for example) can influence the salary. The main statistics of the continuous variables can be found in table 1.

IV. Econometric models and results

We want to examine the marginal impact of having a student on labor income. To do so, we first consider a simple with only the average student loan in households. Moreover, we consider a log-log model since we want to examine how much the student can increase or decrease labor income. This model gives us a first raw idea to test whether our intuition seems plausible. Finally, we obtain that increasing by 1% the average student loan in the household leads to, on average and ceteris paribus, to an increase of 0.064% of labor income and this coefficient is significant at the level 5%. This model gives a first idea but is rather poor, indeed we choose to incorporate the controls we have mentioned before (the parents' education of the

¹ Because we aggregate at the household level, in households with two persons, the informations concerning the sex, the age, the parents education, the major, the highest level education are from the referent.

referee, the highest degree of the referee, sex and age of the referee, the number of children in the household and the major of the referee). When we add these controls, we obtain that an additional 1% of the average student loan in the household leads to, on average and ceteris paribus, to an increase of 0.034% of labor income and this coefficient is significant at the level 5%.² We obtain a lower impact of the student loan because in the previous model, the variable student loan caught all the variations of labor income whereas in this model, these variations can also be caught by all the controls, reducing the marginal impact of the average student loan.

Robust OLS regression with controls :

$$LabInc_i = \tilde{\alpha} + \tilde{\beta}StuLoan_i + \tilde{\gamma}X_i + \epsilon_i$$

3

After having evaluated these two first models, we will now test the assumption of homoskedasticity. Rejecting this hypothesis indicates that the variance of the error term depends on the value of the average student loan and the controls and for instance, a higher average student loan is associated with a higher variance of residuals. To test this assumption, we make a White's test on both previous models and in these two cases, we obtain p-values that are strictly less than 5%, leading to the rejection of homoskedasticity and then our previous models are wrong. To tackle the presence of heteroskedasticity, we now consider two new robust models, one without control and one with controls. In the first robust model without controls, we estimate that an increase of 1% of the student loan per head leads to an increase of 0.028% of labor income and this coefficient is significant at the 5% level. Let us recall that in the non-robust model without controls, the impact was 0.064%, which is more than two times higher. Now we consider the second robust model with the controls and in this model,

² In the annexe 1, we determined that the variable that has the most impact on the coefficient of the student loan is, in decreasing order, the level of education, the sex of the referent, the major of the referent and the level of education of the parents.

³ This is a representation of our econometric model, X_i represents the set of all the controls.

an increase of 1% of the student loan per head leads to an increase of 0.0111% of labor income and this coefficient is significant at the 5% level. In the non-robust model with controls, this effect was 0.034%, which is three times higher.⁴

Regarding the controls, we notice that in the complete model, the education of the father is highly significant whereas the education of the mother is not. *Ceteris paribus*, if the father of the referent finished high school the average labor income in the household rose by 30%. *Ceteris paribus*, a household of only one female, has a labor income of 44% below the labor income of a household with at least one man. *Ceteris paribus*, when the referent doesn't have a college degree, the labor income decreases by 50%.

V. Conclusion and limits

In this paper, we tried to explain whether contracting a student loan has a positive or negative impact on an individual's future labor income in the US. In this matter, we ran many econometric models and came to the following conclusion. For people aged between 25 and 35 years old in the US, an increase of the student loan of 1% yields an increase in the labor income of 0,01%.

We mainly differentiate from the other studies by only considering a sample of individuals aged between 25 and 35 years old, in order to focus on the income at the beginning of an individual's career, while most literature chooses a time frame of the whole career. Furthermore, contrary to most authors that analyzed the impact of student loans, we took into account the labor income. This enabled us to have a more precise analysis of the effect of

⁴ In annexe 2, we made quantile regression of different orders. The impact of the student loan varies depending on which quantile the labor income is situated in. For the 10% lowest labor income the student loan has the highest impact with a coefficient of 0,04. Then the impact gradually declines with the quantile order. It has a negative impact for incomes in the 90th quantile.

having a student loan on an individual's career. By restraining our analysis to the labor income and to individuals in their early career, we exclude some exogenous shocks specific to some individuals that may disturb the result. The individual shocks may be receiving a heritage, job opportunities, or even economic shocks among a specific sector. Although it can depollute the result from specific exogenous shocks, it's important to recall that this choice may affect the validity or relevance of our result. Since individuals take into account the income perceived during their whole career when making a student loan and not the ones perceived during their 10 first career years. Nonetheless, this paper is a good complement to the available literature on the subject by strengthening the conclusion made when taking the whole career into consideration. At last, we take notice that econometric analysis doesn't take into account the individual benefit of pursuing the career path of their choice, by taking the student loan. On the other hand, it doesn't include the burden and stress one may endure due to the student loan.

However, many technical issues were not or hardly treated in our paper. First, despite the use of control variables, the endogeneity problems were not fully mitigated. Indeed, it would have been better to use a 2SLS regression with an instrumental variable for the variable *Student Loan*. One potential candidate would have been the parents' wage. The parents' wage (negatively) impacts the amount of the student loan an individual will contract, so the variables *Student Loan* and *Parents' Wage* are correlated and the relevance condition is hence satisfied. Moreover, it is plausible to think that the parents' wage impacts an individual's future labor income only through the amount of student loan, so that the variables *Labor Income* and *Parents' Wage* are uncorrelated (exogeneity condition). Another problem of our study is induced by aggregating at the household level and taking the average for the variables *Labor Income* and *Student Loan*. By Doing so, we create more variance in the result and omit certain specific cases: for instance, the referent could have a high labor income without having contracted a student loan beforehand, while the spouse could be in the opposite situation.

Aggregating at the level of the household prevents us from taking into account these situations. At last, but not least, we can question the quality of some of our control variables. The variable *Parents' Education* for example does not perfectly measure their real knowledge, the social capital of the family and the environment in which an individual will grow and which will impact its future labor income. A considerable improvement on what we initialized is to determine the optimal amount of student loan depending on the level of education and the major the individual is considering.

One of the main motivations of our subject is the trade off to which is confronted a student in the US, between contracting a loan and aiming for a higher salary but having a debt, and avoiding any loans but targeting a potentially lower wage. We can ask if the choice of field of study is different when the burden of making this loan is assisted by the government.

VI. Tables and Figures

	age of the referent	children	logarithme of the income per head	labor income per head	logarithme of the student loan per head	student loan per head	number of person in the household
Min.	25.00	0.00	0.00	0.00	0.00	0.00	1.00
1st Qu.	28.00	0.00	9.39	12000.00	0.00	0.00	1.00
Median	30.00	1.00	10.13	25000.00	0.00	0.00	2.00
Mean	30.23	1.12	9.24	30138.41	3.72	12334.74	2.71
3rd Qu.	33.00	2.00	10.61	40500.00	9.21	10000.00	4.00
Max.	35.00	9.00	13.11	495000.00	13.46	700000.00	16.00

Table 1: Statistics for the continuous variables

	Model 1	Model 2
(Intercept)	9.86*** (0.03)	8.73*** (0.23)
log_stu_loan_per_head	0.04*** (0.00)	0.01** (0.00)
highst_degree_refMaster		0.02 (0.10)
highst_degree_refNoCollege		-0.50*** (0.06)
highst_degree_refOther		-0.18* (0.08)
highst_degree_refPhD		-0.01 (0.29)
nb_child		-0.06*** (0.02)
majorengineering		0.10 (0.15)
majorother		-0.25** (0.09)
majortechnological		0.09 (0.19)
edu_fa_refcollege		0.43*** (0.08)
edu_fa_refhigh school		0.30*** (0.06)
edu_mo_refcollege		0.14 (0.08)
edu_mo_refhigh school		0.10 (0.07)
age_ref		0.05*** (0.01)
sex_ref		-0.44*** (0.04)
Num. obs.	2427	1918

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table 2: Results for the robust regressions

5

⁵ We can notice that upon the different models, the number of observations varies. This is due to omitted responses in the survey.

VII. References

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L. Violante. ‘Education Policy and Intergenerational Transfers in Equilibrium’. *Journal of Political Economy* 127, no. 6 (December 2019): 2569–2624. <https://doi.org/10.1086/70224>

Abel, Jason R., and Richard Deitz. ‘Do the Benefits of College Still Outweigh the Costs?’ SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 1 August 2014. <https://papers.ssrn.com/abstract=2477864>

Avery, Christopher, and Sarah Turner. “Student Loans: Do College Students Borrow Too Much—Or Not Enough?” *The Journal of Economic Perspectives* 26, no. 1 (2012): 165–92. <http://www.jstor.org/stable/41348811>

Barr, Andrew, Kelli A. Bird, and Benjamin L. Castleman. ‘The Effect of Reduced Student Loan Borrowing on Academic Performance and Default: Evidence from a Loan Counseling Experiment’. *Journal of Public Economics* 202 (1 October 2021): 104493. <https://doi.org/10.1016/j.jpubeco.2021.104493>

Behrman, Jere R., et Paul Taubman. « The Intergenerational Correlation Between Children’s Adult Earnings and Their Parents’ Income: Results from the Michigan Panel Survey of Income Dynamics ». *Review of Income & Wealth* 36, n° 2 (juin 1990): 115-27. <https://doi.org/10.1111/j.1475-4991.1990.tb00275.x>

Britton, Jack, Laura van der Erve, et Tim Higgins. « Income Contingent Student Loan Design: Lessons from around the World ». *Economics of Education Review, Higher Education Financing: Student Loans*, 71 (1 août 2019): 65-82. <https://doi.org/10.1016/j.econedurev.2018.06.001>

Daniel H. Cooper and J. Christina Wang. ‘Student Loan Debt and Economic Outcomes’. Boston, Federal Reserve Bank of.. Federal Reserve Bank of Boston, 2 December 2014. <https://www.bostonfed.org/publications/current-policy-perspectives/2014/student-loan-debt-and-economic-outcomes.aspx>

Elliott, William, and Melinda Lewis. ‘Student Debt Effects on Financial Well-Being: Research and Policy Implications’. *Journal of Economic Surveys* 29, no. 4 (September 2015): 614–36. <https://doi.org/10.1111/joes.12124>

Elliott, W., & Nam, I. (2013). *Is Student Debt Jeopardizing the Long-Term Financial Health of U.S. Households?* St. Louis, MO: St. Louis Federal Reserve Bank. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.401.2486&rep=rep1&type=pdf>

Higgins, Tim, et Mathias Sinning. « Modeling Income Dynamics for Public Policy Design: An Application to Income Contingent Student Loans ». *Economics of Education Review* 37 (1 décembre 2013): 273-85. <https://doi.org/10.1016/j.econedurev.2013.08.009>

Houle, Jason N. « Disparities in Debt: Parents' Socioeconomic Resources and Young Adult Student Loan Debt ». *Sociology of Education* 87, n° 1 (2014): 53-69.

Keane, Michael P., and Kenneth I. Wolpin. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review* 42, no. 4 (2001): 1051–1103. <https://www.jstor.org/stable/826985>

Meta Brown and Sydnee Caldwell. 'Young Student Loan Borrowers Retreat from Housing and Auto Markets', 17 April 2013. <https://libertystreeteconomics.newyorkfed.org/2013/04/young-student-loan-borrowers-retreat-from-housing-and-auto-markets/>

Rothstein, Jesse, and Cecilia Elena Rouse. 'Constrained after College: Student Loans and Early-Career Occupational Choices'. *Journal of Public Economics* 95, no. 1–2 (February 2011): 149–63. <https://doi.org/10.1016/j.jpubeco.2010.09.015>

Rouse, Cecilia Elena. « Low-Income Students and College Attendance: An Exploration of Income Expectations ». *Social Science Quarterly (Wiley-Blackwell)* 85, n° 5 (15 décembre 2004): 1299-1317. <https://doi.org/10.1111/j.0038-4941.2004.00277.x>

Sakellariou, Chris, et Zheng Fang. « Self-Efficacy and Interest in STEM Subjects as Predictors of the STEM Gender Gap in the US: The Role of Unobserved Heterogeneity ». *International Journal of Educational Research* 109 (1 janvier 2021): 101821. <https://doi.org/10.1016/j.ijer.2021.101821>.

Sirin, Selcuk R. « Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research ». *Review of Educational Research* 75, n° 3 (2005): 417-53.

Weichselbaumer, Doris, et Rudolf Winter-Ebmer. « A Meta-Analysis of the International Gender Wage Gap ». *Journal of Economic Surveys* 19, n° 3 (juillet 2005): 479-511. <https://doi.org/10.1111/j.0950-0804.2005.00256.x>.